Introduction *to* Transformational Grammar

Kyle Johnson
University of Massachusetts

Fall 2007

Contents

Preface

These are the always evolving notes from an introductory course on syntactic theory taught at the University of Massachusetts at Amherst. Its target audience is first-year graduate students, but no background exposure to syntax is presupposed. The course is taught together with a matching introductory course on semantics which, in the Fall of 2007, is led by Angelika Kratzer. These notes augment a set of readings, which are:

> Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Massachusetts: M.I.T. Press, Chapter 1.
>
> Stowell, Timothy. 1981. Origins of phrase structure. Doctoral Dissertation, Massachusetts Institute of Technology, Chapters 1–3.
>
> Pollock, Jean-Yves. 1989. Verb movement, UG and the structure of IP. *Linguistic Inquiry* 20:365–424.
>
> Vikner, Sten. 1995. *Verb movement and expletive subjects in the Germanic languages*. Oxford: Oxford University Press, Chapter 3.
>
> Larson, Richard. 1988. On the double object construction. *Linguistic Inquiry* 19:335–392.
>
> Johnson, Kyle. 1991. Object positions. *Natural Language and Linguistic Theory* 9:577–636.
>
> Kratzer, Angelika. 1996. Severing the external argument from its verb. In *Phrase structure and the lexicon*, ed. Johan Rooryck and Laurie Zaring, 109–137. Dordrecht, The Netherlands: Kluwer Academic Publishers.
>
> Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht, The Netherlands: Foris Publications, Chapter 3.

Chomsky, Noam. 1986. *Knowledge of language*. New York, New York: Praeger Publishers, Chapter 3.

Reinhart, Tanya, and Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24:657–720.

Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, Massachusetts: M.I.T. Press, Chapters 1–5.

Chomsky, Noam. 1995. Bare phrase structure. In *Government binding theory and the minimalist program*, ed. Gert Webelhuth, 383–439. Oxford: Oxford University Press.

Fox, Danny, and David Pesetsky. 2004. Cyclic linearization of syntactic structure. *Theoretical Linguistics* 31:1–46.

1

The Subject

Linguistic theory, and so syntactic theory, has been very heavily influenced by learnability considerations in the last fifty-some years, thanks largely to the writings of Noam Chomsky. If we decide that syntactic theory is charged with the duty of modeling our knowledge of language, then we can make some initial deductions about what this knowledge, and therefore our model of it, should look like from some simple observations. This knowledge must interact efficiently with the cognitive mechanisms involved in producing and comprehending speech, for instance. It must also be acquirable by any normal infant exposed to speakers of the language over six or so years. A number of considerations combine to make the task of acquiring knowledge of a language look very difficult indeed: the complexity of the acquired grammar, the amount of information that needs to be acquired, the attenuated nature of the information available to the child, etc. It is made even more puzzling by the fact that children appear to complete this task with relative ease in a comparatively short period of time and that the course of acquisition appears to go through a set schedule of stages. There is clearly a problem: If languages are as complex as we think they are, then how can these impossibly complex objects possibly be learned?

1.1 Linguistics as learning theory

It is Chomsky's proposal that Syntactic Theory itself should contribute to solving this dilemma. The classical formulation of this idea (see *Aspects* and *The Sound Pattern of English*) characterizes the situation as follows. Think of a grammar of L (G_L) (this is what Chomsky (1986b) calls "I-Language") as a set of rules that generates structural descriptions of the strings of the language L (Chomsky (1986b)'s E-language). Our model of this grammar is descriptively adequate if it assigns the same structural descriptions to the strings of L that G_L does. We can think of the learning process as involving a selection from the Universe of G_Ls the very one that generates these structured strings of the L to be acquired.

The learning problem can now be stated in the following terms: how is it that the learning procedure is able to find G_L when the universe of Gs is so huge and the evidence steering the device so meager.

One step towards solving this problem would be to hypothesize that the universe of Gs has a structure that enables convergence on G_L given the sort of information that the child is exposed to. This is Chomsky's proposal. It amounts to the claim that there are features of Gs which are built-in: certain properties which distinguish the natural class of Gs from the rest. There is a kind of meta-grammar of the Gs, then, which is sometimes referred to with the label Universal Grammar. Chomsky further hypothesizes that these properties are biologically given: that it is something about the construction of the human brain/mind that is responsible for the fact that the class of Gs are the way they are. This argument, the one that leads from the observation that G_Ls have features that are too complex to be learned to the conclusion that the universe of Gs is constrained is often called "The Poverty of the Stimulus" argument. It is a classic from Epistemology, imported with specific force by Chomsky into linguistics.

This way of setting up the problem, note, allows for the Universe of Gs to be larger than the learnable Gs. There could, for instance, be constraints imposed by the parsing and production procedures which limit the set of Gs that can be attained. And it's conceivable that there are properties of the learning procedure itself — properties that are independent of the structure of Gs imposed by Universal Grammar — that could place a limit on the learnable Gs. Universal Grammar places an outside bound on the learnable grammars, but it needn't be solely responsible for fitting the actual outlines of that boundary. It's therefore a little misleading to say that the set of "learnable Gs" are those

characterized by Universal Grammar, since there may be these other factors involved in determining whether a grammar is learnable or not. I should probably say that Universal Grammar carves out the "available Gs," or something similar. But I will instead be misleading, and describe Universal Grammar as fixing the set of learnable Gs, always leaving tacit that this is just grammar's contribution to the learnability question.

Chomsky proposes, then, that a goal of syntactic theory should be to contribute towards structuring the universe of Gs. He makes some specific proposals about how to envision this in *Aspects of The Theory of Syntax*. He suggests that syntactic theory should include an evaluation metric which "ranks" Gs. A syntactic theory that has this feature he calls explanatory. Thus "explanatory theory" has a specific, technical, sense in linguistic theory. A theory is explanatory if and only if it encapsulates the features that ranks Gs in such a way that it contributes to the learnability problem, distinguish the learnable Gs from the unlearnable ones. This criterion can help the syntactician decide whether the model of G_L he or she has proposed corresponds exactly to G_L. In particular, the many descriptively adequate models of G_L can be distinguished on this basis: we should select only those that are ranked highly by the evaluation metric. These grammars meet the criterion of explanatory adequacy.

A very important role, therefore, is played by the evaluation metric. At the time of *Aspects*, the learning procedure was conceived of as a process very much like that which the linguist goes through. The child builds a battery of rules which generate the strings of L. The evaluation metric steering this process was thought to have essentially two parts: a simplicity metric, which guides the procedure in its search through the space of grammars, and inviolable constraints, which partitions the set of Gs into the learnable ones and the unlearnable ones. Thus, for example, we might imagine that rules which used fewer symbols could be defined as "simpler" than ones that used a greater number of symbols. Inviolable constraints might be those, for example, expressed as part of the principles which place constraints on the way that strings can be partitioned into groups, and therefore simply removes from the universe of Gs a great many possible Gs. Let's call these models of Gs "rule based," because the simplicity metric is defined as a procedure that constructs rules, and the companion picture of the acquisition process the "Little Linguist" model.

To take a concrete example, imagine that the principles which limit how words are strung into groups — one particular version of which goes by the name "$\overline{\text{X}}$ Theory" — imposes the following constraints.

$$XP \rightarrow \{ (ZP), \overline{X} \}$$
$$\overline{X} \rightarrow \{ \overline{X}, (YP) \}$$
$$\overline{X} \rightarrow \{ X^0, (WP) \}$$

Understand "$\{\alpha, \beta\}$" to signify that α and β are sisters, and "(α)" to indicate that α is optional. Let W, X, Y and Z range over kinds of lexical items (e.g., "noun," "verb," "preposition," and so on). And, finally, let "→" mean: "consists of." The groups here, known as phrases, are the XP and $\overline{X}$ in the formulas. These constraints, then, leave to the learner only the matter of filling in the variables W, X, Y and Z, and discovering their linear order. As the child goes from step to step in matching the grammar he or she is constructing with the information coming in, these are the only decisions that have to be made. If we imagine that this set of options were to be operationalized into a concrete decision tree, then we could see this as constituting a kind of "simplicity metric." It would constitute a procedure for searching through the space of learnable grammars that imposes an order on the grammars, enabling a deterministic method for converging at a particular grammar when exposed to a particular linguistic environment. Additionally, $\overline{X}$ Theory provides an absolute cap on the possible phrases and, in this respect, constitutes an inviolable constraint as well. If every language learner is equipped with this $\overline{X}$ Theory, then they will converge more or less on the same G_L when presented with the information that being in the environment of speakers of L provides. If there are differences in the G_L's that learners converge on, these will trace back to different decisions these learners have made about the identity of W, X, Y and Z, or how their linear order is determined. If the rest of a model that incorporates these constraints is correct, then, it should allow any language learner to pick out a G_L very close to the G_L giving shape to the speech in that learner's environment.

Let's consider another example involving transformational rules, one that Chomsky often points to. Transformational rules map one syntactic representation, D-structure, to another, S-structure, typically by way of moving constituents. Interestingly, it appears that all such rules are "structure dependent." That is, they make reference to the relative structural positions of the moved thing and the position it is moved to. They don't, for example, make reference to points in a string on the basis of their position relative to some numerical count of formatives. Thus "Wh-Movement" moves maximal projections that meet certain criteria to particular positions in a phrase marker. And this op-

eration is governed by a set of constraints that make reference to the relation between these points solely in terms of structure. There is no rule, for example, like Wh-Movement but which affects terms based on how far apart they are numerically. Thus, the learning procedure will never have to entertain the hypothesis that G_L should contain such rules.

In both cases, the classic argument for distinguishing the inviolable constraint from the simplicity metric follows very closely the logic of the poverty of stimulus argument. Because it is difficult to see (maybe even provably impossible) how such things as $\overline{X}$ Theory or structure dependence could be learned, they must belong to the features that define the universe of Gs. And because they are overarching properties of the rules in some G_L, they also have the right form to be inviolable constraints.

There is another argument towards the same end which has gained increasing influence in the last couple decades; and this one comes to us through the narrowly linguistic study of language typology, and only tangentially from learnability considerations. I will call it "Humboldt's argument," though it no doubt has an earlier champion. Humboldt's argument is based on the observation that there are certain properties that appear to hold true of all G_Ls. This can be explained, Humboldt argues, only if the universe of Gs is constrained to just those which have the relevant, universal, properties. Like Chomsky, Humboldt relates this to the construction of the mind, and uses the language of learnability in his account. He puts it this way:

> Since the natural inclination to language is universal to man, and since all men must carry the key to the understanding of all languages in their minds, it follows automatically that the form of all languages must be fundamentally identical and must always achieve a common objective. The variety among languages can lie only in the media and the limits permitted the attainment of the objective.
>
> (von Humboldt (1836))

(One might read the last sentence of this passage as making the distinction, touched on above, between aspects of Universal Grammar ("the media") and the limits our cognition places on exploiting UG ("the limits permitted the attainment of the objective").) So, like Chomsky, he supposes that there is a Universal Grammar, a feature of the mind, which constrains the form that languages may have. But his perspective is different from Chomsky's. He expresses the notion of Universal Grammar not in terms of learning theory, or through

the glass of the Poverty of the Stimulus argument, but from the perspective of language variability. He links limits on language variability to a universal ability he sees in human psychology to acquire a language.

Humboldt's goal is an explanation for the observed limits in variability of the grammars of extant languages. One might imagine that there are explanations for these limits that do not involve, as Humboldt proposes, constraints imposed by human psychology. Similarities in extant languages might reflect their common ancestry: if all languages descend from a common one, then features that are shared among them could simply be vestiges of the ancestral language that historical change has left untouched. This is the thesis of monogenesis. I think it's possible to read Sapir as advancing this alternative. Sapir is commonly associated with the position exactly opposite to Humboldt's; in Sapir's words:

> Speech is a human activity that varies without assignable limit as we pass from social group to social group, because it is a purely historical heritage of the group, the product of long-continued social usage.
>
> (Sapir 1921, p. 4)

But, perhaps because of his vagueness, it's possible to credit Sapir with a more sophisticated view. One that assigns the universal properties of languages to the detritus of historical change:

> For it must be obvious to any one who has thought about the question at all or who has felt something of the spirit of a foreign language that there is such a thing as a basic plan, a certain cut, to each language. . . . Moreover, the historical study of language has proven to us beyond all doubt that a language changes not only gradually but consistently, that it moves unconsciously from one type towards another, and that analogous trends are observable in remote quarters of the globe.
>
> (Sapir 1921, pp. 120-121)

Perhaps the common properties of extant (and known) languages are a function of two facts: all languages descend from a common language, and the forces that cause languages to change are not fully random — they preserve certain features and change others only according to some "basic plan." If historical relatedness is to explain the common traits that extant languages have,

some limit must be placed on how languages change and diverge. Otherwise, language change would act as a kind of randomizer that, over time, would destroy the limits in variability that we observe. Mongenesis needs to be coupled, then, with a theory of diachrony that characterizes the limits it imposes on change. Could it be, then, that the similarities in languages are all due to these laws of diachrony?

This seems to me to be a coherent account for language variability. But it may be just a disguised version of the Chomsky/Humboldt hypothesis that the limits of human cognition are responsible for the constraints on linguistic variation. The thesis of monogenesis entails that language variation is solely the product of historical change, as Sapir's quotes makes clear. So we expect that languages vary in features which historical change can affect, but will remain similar in those ways that are immutable. Which of the features appear as language universals, then, is determined by the internal mechanisms of historical change, and the limits thereon. What are the internal mechanisms of historical change? The only proposal I know of is that historical change is a by-product of language acquisition. It is the accumulation of the small mismatches in G_Ls that successive generations of language acquirers select. Language acquisition, the poverty of the stimulus argument tells us, is guided by Universal Grammar. So even granting the diachronic argument for language universals, we see that as historical change weeds out the mutable properties from the immutable ones, the properties it leaves are those that characterize Universal Grammar. The antidote for the argument I have blamed on Sapir, then, involves bringing the poverty of the stimulus argument into play. I don't know if Humboldt's argument can stand against this alternative unaided.

But even if it can't, it provides us with another way of viewing how to factor out the components of the evaluation metric. Following the logic of Humboldt's argument, what we expect is that language comparison should give us a means of separating inviolable constraints from the evaluation metric. The inviolable constraints will be (among) those things found in all languages; the differences in languages are to be credited to the evaluation metric. Put somewhat differently, an explanatory theory is to give us both how languages cannot be constructed, and how their construction can vary. The data it must fit, then, emerges only once languages are compared: for not only does this allow the universals to be clearly discerned, but it is only through this means that the particulars of language variation are known.

When this method of factoring out the universals in Gs is followed in earnest, a rather different picture of various G_Ls emerges; and a very different concep-

tion of the language acquisition procedure becomes available. This course is meant to illustrate these emerging pictures in detail.

1.2 The evidential basis of syntactic theory

If linguistics is one part of the study of human cognition, in the sense just described, then syntax can be described as that subdiscipline of linguistics which seeks to discover what speakers know about how to arrange the words of their language into meaningful sentences. Because speakers are not conscious of the principles that characterize this knowledge, the syntactician must make recourse to indirect means of determining these principles. The syntactician's first task, then, is to determine how to find evidence that reflects the nature of this knowledge.

One plausible source of relevant information comes from observing how speakers put this knowledge to use. We could, for instance, collect the utterances from some speaker and look for generalizations in these utterances from which evidence about the underlying knowledge-base can be gleaned. This is rarely done, however, as there are few instances of such collections that arise naturally, and to assemble them from scratch is onerous enough to have been avoided. With the exception of studies of prodigious literary figures, there are vanishingly few attempts at linguistic studies that go this route.

More common is to study the linguistic utterances of a group of speakers. This is standardly done by using the dictionary maker's device of combing texts and newspapers for examples. There are several excellent "parsed" corpora of this sort,[1] and even corpora of spoken utterances[2] can be found. With the advent of the World Wide Web, it has become possible to search a very large collection of sentences, and more and more linguists are availing themselves of this resource. This technique has the unique advantage of allowing one to determine frequencies as well. It is possible, for example, to judge how rare some particular arrangement of words is relative to some other, or to find statistically significant correlations between, say, the position of an argument relative to its predicate and the person or number marked on that argument. Some linguistic theories are specifically designed to model these sorts of frequency data.[3]

1 See Marcus, Santorini, and Marcinkiewicz (1993), for example.
2 See Godfrey, Holliman, and McDaniel (1992).
3 See the papers in Bod, Hay, and Jannedy (2003) for some recent examples of statistically based

There are some serious pitfalls to using group corpora, however. One is simply that it obliterates differences among speakers and treats the data as if it were all manufactured by the same grammatical system. Since nothing is known about the producers of these sentences – they may include speakers of different dialects and speakers for whom the language in question is non-native or has been influenced by another language, for instance – this could be a serious source of error. Without some measure of the heterogeneity of the speakers who produced the corpus, it is very difficult to judge how faithfully it represents the syntactic knowledge of any one of those speakers.

Another shortcoming is that linguistic behavior, even of one individual, is not a faithful projection of the knowledge that that individual has of his or her language. People say sentences whose syntactic form is at odds with what they would otherwise deem well-formed. A significant proportion of any corpus could be made up of such "mistakes," and indeed it would be prudent to assume so, given the degree to which misshapen sentences populate the utterances of such well-placed contributors to corpora as George W. Bush. There is a distinction between a speaker's linguistic "performance" and his or her linguistic "competence," to use the names Chomsky gives to this distinction. Corpora level this distinction.

For these reasons, then, group corpora contain an unknown amount of data that should be weeded out. They contain examples of sentences that are produced by speakers whose grammatical systems differ, and they contain sentences that are not representative of any grammatical system. But group corpora are not only noisy with error, they are also mute about certain kinds of information.

One important piece of evidence that corpora cannot provide concerns where speakers draw the line between impossible and possible forms in their language. This distinction is easiest to elicit in linguistic domains where there are a comparatively small number of relevant forms. For example, the morphological and phonological inventories of any one speaker at any one time is reasonably small and it is therefore salient when a novel morphological or phonological form is introduced. For many such novel forms, speakers are capable of distinguishing those that are admissible members to their languages and those that are not. Most English speakers I have asked, for instance, can tell that *blick* ([blɪk]) is an admissible addition to their lexicon but that *bnick*

corpora studies, and the work of Paul Boersma (e.g., Boersma and Hayes (2001)) for a theory that is designed to model statistical data of this sort.

([bnɪk]) is not. Presumably this ability to distinguish admissible from inadmissable forms is due to the knowledge speakers have of their language, and so it is an important piece of information about how that knowledge is constituted. A typical way of characterizing this distinction goes as follows. The phonology of a language permits many forms that are not exploited by the lexicon of that language (e.g., [blɪk]). Which of these forms are used and which are not is completely extragrammatical. By contrast, because the phonology of a language limits the forms that are available to that language (e.g., English prevents the onset cluster [bn]) these forms (e.g., [bnɪk] in English) will be blocked from its lexicon. The absence of these forms is determined by the grammar; they are said to be "ungrammatical," and when they are cited, they are prefixed with the diacritic "*" to indicate their status.

The same distinction can be elicited for sentences, although because of the larger number of forms involved it is more difficult to recognize a novel sentence. Consider, by way of illustration, the pair of sentences in (1).

(1) a. Whenever the earth revolves around its equator, the moon begins to rotate about its axis.
    b. Whenever the earth revolves around its equator, the moon begins itself to rotate about its axis.

I judge (1b) to be an impossible English sentence, and (1a) to be a possible one. Because I read very little science fiction, I think it's likely that both sentences are novel for me, but I do not have the certainty about this that I have about *blick* and *bnick*. I recognize that there are considerably more sentences that I have encountered than there are words I've encountered, and consequently I also recognize that it is likelier that I will mistake a sentence as novel than it is that I will mistake a word as novel. Nonetheless, most linguists would agree that the contrast in (1) is of the same kind that distinguishes *blick* from *bnick*. It does seem unlikely that the distinction could be reduced to one of novelty. After all, I am roughly as certain of the novelty of (1a) as I am of the novelty of (1b) and yet this does not affect the strength of my judgement concerning their Englishness. It seems probable that my ability to judge the difference between (1a) and (1b) traces back to an ability my syntactic knowledge gives me to judge well-formedness.

This distinction between grammatical and ungrammatical forms is important because it seems to tap directly into a speaker's linguistic knowledge. Studying corpora cannot provide what is needed to see this distinction; corpora conflate ungrammatical and grammatical but non-occurring forms. For

this reason, and because of its noisiness, I will not use data from corpora in these lectures. But do not forget that corpus studies, and so far as I know only corpus studies, can provide statistical data, for this might be an important resource in forming a complete model.

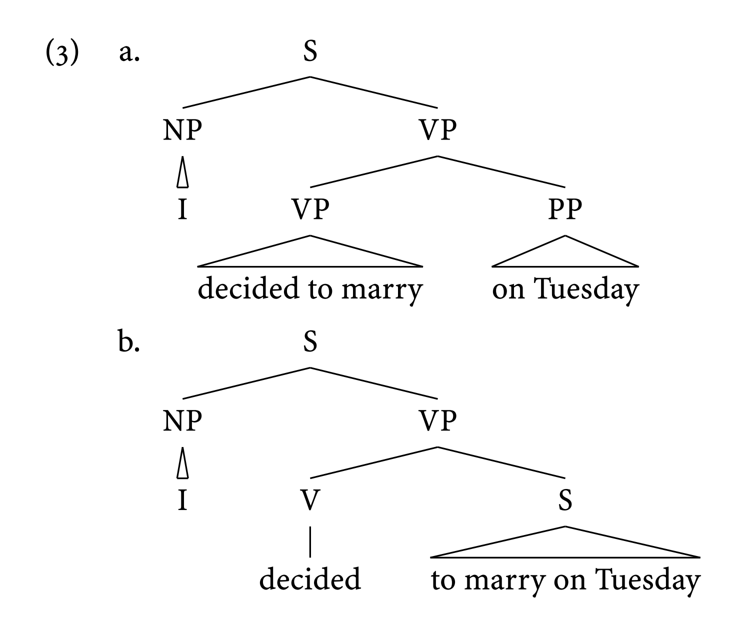
Instead, the central piece of evidence used in these lectures will be elicited grammaticality judgments. This has become the standard tool for syntactic analysis, and much of the literature relies on it. Elicited grammaticality judgments have their own shortcomings. There are special problems attendant with grammaticality judgments of sentences. Because sentences are very complex objects, and are frequently longer than the small memory buffer that our online processors are equipped with, there are failures of sentence processing that might easily be mistaken for judgments of ill-formedness. A famous example meant to be illustrative of this distinction comes from strings that are ambiguous with respect to the placement of some late occurring phrase. The pair of sentences in (2) illustrates.

(2) a. I decided to marry on Tuesday.
b. I decided that my daughter should marry on Tuesday.

Upon reflection, most speakers will recognize that (2a) has two meanings. It can assert that the time of my decision to marry was Tuesday, or it can assert that what my decision was was to marry on Tuesday. As we will see, this ambiguity reflects the fact that (2) maps onto two sentences, whose difference in syntactic structure is responsible for the two meanings. The first meaning corresponds to a structure which groups the words as sketched in (3a) on the following page, whereas the second interpretation corresponds to the syntactic structure shown in (3b).

Unlike (2a), (2b) seems to have only the second of these two meanings. It can assert that my decision was for my daughter to marry on Tuesday, but it does not seem to say that the time of my decision was Tuesday. At present, this difference in (2a) and (2b) is thought to be due to constraints of sentence processing, and not the well-formedness conditions of sentences. The relevant difference between these examples is the number of formatives between the word *decided* and the prepositional phrase *on Tuesday*. As that number grows beyond what can be held in working memory, the processor is forced to start making decisions about how to parse the initial portions of the string. These decisions favor a parse in which later material is made part of more deeply embedded phrases. Thus, in the case of (2b) it favors the structure in (4b) over that in (4a) on the next page. On this account, then, it is not that there is a

(3)



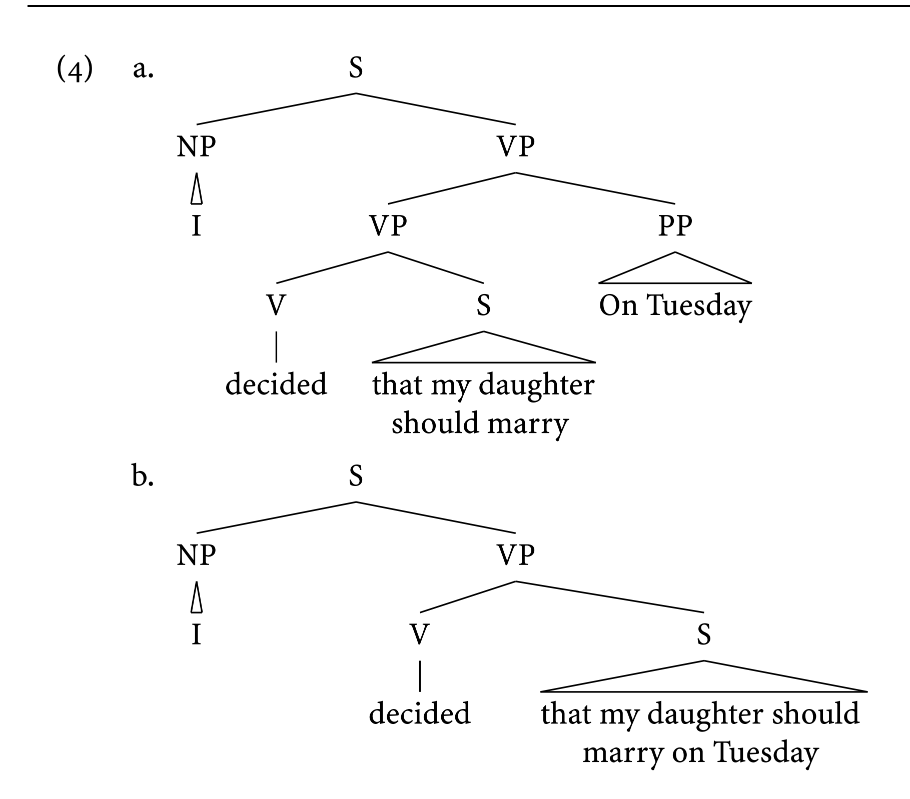
difference in the syntactic well-formedness conditions which causes speakers' differing judgments about (2a) and (2b). Instead, because of the relative difficulty that (2b) presents to the on-line processor, one of the syntactic representations associated with this string (i.e., (4a)) becomes difficult to perceive. This effect of the on-line processor is what Kimball called "right association."[4]

In general, judgments of well-formedness will not be able to distinguish those sentences that do not conform to the constraints of the grammar from those that do conform to those constraints but present problems for the on-line processor.[5] There is no simple way of distinguishing these cases; they can be separated only through analysis. In the case of (2), the decision that the effect is not grammatical but, instead, the result of the processor comes partly from finding no good grammatical way of distinguishing the cases and partly from finding that manipulating factors relevant for the processor determines whether the effect materializes.

Another similar difficulty involves the fact that the meanings which sentences convey are typically bound to the context of a larger discourse. Inevitably, then, grammaticality judgments are going to be confounded with whether or not there is a discourse in which that sentence could function. Suppose, for

4 See Kimball (1973), Frazier (1978) and Gibson (1998).
5 Chomsky and Miller (1963) is an early, and still useful, examination of this distinction.

(4) a.



instance, that you are trying to determine the distribution of a process called "VP Ellipsis," which allows a sentence to go without a normally required verb phrase. VP Ellipsis is responsible for allowing the bracketed sentence in (5) to go without a verb phrase in the position marked "Δ."

(5) Jerry annoyed everyone that [$_S$ Sean did Δ].

If you expose English speakers to the examples of VP Ellipsis in (6), you may find that they judge them ungrammatical.

(6) a. * Whomever she did Δ got better.
b. * Everything for her to Δ was hard.

One might be tempted by these examples to the hypothesis that VP Ellipsis is blocked within subjects. But if the examples in (6) are embedded into an appropriate discourse, English speakers will find (6a) well-formed while (6b) remains ungrammatical.

(7) a. Whomever Sally didn't tutor got worse but whomever she did Δ got better.

(8) Two Laws of Elicitation

a. The sentences for which you elicit a grammaticality judgement should be embedded in a discourse that makes the meaning that sentence would have salient.

b. Every suspected ungrammatical sentence should be part of a minimal pair, the other member of which is grammatical.

b. * Everything for him to do was easy and everything for her to Δ was hard.

The problem with (6a) is that recovering the meaning of the elided VP cannot be done without a larger context, and the grammaticality of sentences with VP Ellipsis in them depends in part on recovering the meaning of the elided VP. There is nothing syntactically ill-formed with the VP Ellipsis in (6a), however, as we see when this context is provided. By contrast, neither the context in (7b) (nor any other that I have found) improves the goodness of (6b). There is something ill-formed about the syntax of this example.

These two problems are similar. In both, the difficulty is in distinguishing judgments of ungrammaticality from other types of ill-formedness. The effect of these difficulties can be lessened if the following two practices are used in eliciting judgments.

First, embed the sentences whose well-formedness you wish to determine in discourse contexts that make the meaning these sentences should have available and salient. This helps remove the second problem.

Second, for every sentence you suspect to be ungrammatical, present your informant with a matching sentence which you suspect to be grammatical. These two sentences – the suspected grammatical and the suspected ungrammatical one – should differ minimally. Your aim should be to remove all differences between these two sentences except for the factor that you suspect is responsible for the ungrammaticality. This will help mitigate processing effects, as the two sentences will end up matched in length and close to matched in complexity. It will also help remove any other confounds which might be responsible for the ungrammaticality of the sentence you wish to test.

These practices are rarely used, unfortunately. As a result, the history of syntactic research is littered with dead ends and wrong turns that have resulted from errors in the empirical base. Don't fall victim to these errors. Wherever you can, follow the Two Laws of Elicitation. In these lectures, I will

sometimes violate (8a) whenever I haven't found a context that improves an ill-formedness judgement. In these cases, my guess is that the ungrammaticality of the sentence is not tied to its information content. Similarly, I will occasionally fail to give a minimal pair when I feel that the ungrammaticality of the sentence involved is dramatic enough to be obvious. For instance, examples such as (9) are so clearly violations of English sentence structure, that I cannot imagine a discourse context that could improve them, nor would minimally contrasting grammatical examples help remove the possibility of a processing effect.

(9) a. * Many happy the puppies barked.
    b. * She talked people to.
    c. * He ate should apples.

I do this partly because it will make the exposition cleaner, but obviously also because I am lazy. It would be wise to maintain a healthy skepticism about the data I present when I've taken these shortcuts.[6]

There is one last danger in relying on elicited grammaticality judgments, and it is the mundane and familiar one of introducing bias. It is a commonplace among experimental psychologists that eliciting psychological data can involve very subtle ways of introducing bias. Whenever the judgments are less clear than obvious cases like (9), the syntactician should clearly not rely on her or his own judgments. In these cases only judgments elicited from naïve informants will do. And in eliciting those judgments, the syntactician should adopt some of the techniques developed by experimental psychologists. Produce a survey of examples that include the sentences you wish to find judgments for but include irrelevant "fillers" as well. Those sentences should be crafted in accordance with the Two Laws of Elicitation. Then present this survey to a number of speakers native in the relevant language, controlling as best as possible for dialect variation. Finally, present the items in the survey in a randomized order, mitigating any bias that the order of presentation might introduce. When reporting data, you should also report the number of informants you have used, and make a note of any variation in the judgments you have encountered. While these safeguards wouldn't satisfy the rigorous numerical criteria of the experimental psychologist, they will go a long way towards re-

6 I will also violate (8) when I am reporting data from the literature in which (8) have not been followed.

moving error and making the data you report comparable to the data someone else gathers.

Grammaticality judgments, then, will be the central evidence used here in uncovering the principles that constitute a speaker's syntactic knowledge. There is one other kind of datum that is important to the syntactician. As we've seen, the syntax of sentences is intimately tied to the meanings they convey. It is the semanticist's job to discover the principles that allow users of language to extract these meanings. One of the central principles of semantics is the law of compositionality.

(10) The Law of Compositionality
The meaning of a string of words, α, is derivable from the meanings of the words that α contains.

As a result of the Law of Compositionality, there is a regular and productive relationship between the syntactic structure of a sentence and the meaning it conveys. This makes it possible to use the meaning a sentence has to draw conclusions about the syntax of that sentence. This requires the assistance of the semanticist, of course, for an intimate knowledge of the rules of semantic composition are needed to draw these conclusions. In recent years, this source of evidence has grown in importance, and it will be an important component of these lectures.

2

Phrase Structure

Our goal is to model the processes by which arrangements of words are recognized as forming grammatical sentences. As sketched in the previous chapter, this involves discovering how those processes vary from speaker to speaker, for only in this way will we get a handle on what features of these processes are universal and on how they are permitted to vary. I presuppose that readers of these notes have some familiarity with English, and so we'll begin this task by investigating those processes that are responsible for the English speaker's grammaticality judgments.

Our first observation is that we can get very far in this task using very little information about the words involved themselves. A great deal about the processes that determine well-formed arrangements of words can be characterized using nothing more than the morpho-syntactic "category" that the words belong to. This can be appreciated by virtue of the "Novel Form" argument, which is laid out in (1).[1]

(1)
  a. If enough information is introduced with a novel word to enable the individual learning that word to recognize its category, then
  b. The individual knows which arrangements it can grammatically combine in.
  c. Hence, it must be category membership to which these processes refer.

1 I thank Jay Keyser for teaching me this argument.

If I introduce you to the word [blowrsnɪk] (‘bloresnick’), for instance, and tell you that it refers to unwanted facial hair, you will be able to determine that only (2b) of the examples in (2) is grammatical.

(2) a. It bloresnicked.
    b. He removed the long bloresnick.
    c. She finds Sammy bloresnick.
    d. He made his face bloresnick.

There is a close connection between the “category” a word belongs to and its meaning, an issue we will revisit shortly. It is tempting to think that morphosyntactic category could, perhaps, be reduced to something about the meanings of the words involved. At present, however, there is no consensus about how to recast categories into wholly semantic terms and, given the differences in how languages seem to map meanings onto categories, there are real problems to overcome in doing so. It’s also not clear that the pattern of grammaticality judgments English speakers give to (2) rests on differences in how their meanings are arrived at. Why can’t the meanings of *it* and *bloresnick* in (2a) combine in the same way that they seem to in (3a), for instance? And why can’t (2c) get a meaning like that assigned to (3b)?

(3) a. It was bloresnick.
    b. She finds Sammy to be bloresnick.

There is information, then, about the kinds, or types, of words that is not identical to the kinds, or types, of things they refer to, and this information seems to underlie our knowledge about grammatical combinations of words.

Let’s see how far we get in modeling grammaticality judgments just by paying attention to morphosyntactic category.

2.1 Substitution Classes

Our lexicon of words, then, is partitioned into sets — categories — and some of our knowledge about which groups of words are grammatical is based on membership in these sets. We can use the traditional grammarian’s terminology for these sets. *Bloresnick*, for instance, is a noun; *find* is a verb; *long* is an adjective, and so on. A string that is made up of a noun followed by a verb followed by a noun is judged to be a grammatical sentence in English (witness (3a)) whereas a string made up of two adjacent nouns is not (compare (2a)). If

we were to look at a large assortment of strings, we would be able to discover patterns that distinguish the grammatical sentences from the ungrammatical sentences. We would discover that categories have a particular distribution in the grammatical strings.

Zellig Harris argued that morpho-syntactic category should be defined in just these terms.[2] Specifically, "noun," "verb" and so on are "substitution classes" of vocabulary items. They are substitution classes in the sense that there is a set of positions within a sentence into which any member of that class can be substituted preserving the grammaticality of the sentence. For instance, any word that can be grammatically placed in the spot marked with "___" in (4) falls within the subset of vocabulary items we know as "nouns."

(4) the ___ exists

This is indicated by considering the lists of sentences in (5)-(10).

(5) The lamp exists.
The girl exists.
The sky exists.
The streetcar exists.
⋮

(6) * The happy exists.
* The blue exists.
* The short exists.
* The flat exists.
⋮

(7) * The in exists.
* The out exists.
* The from exists.
* The on exists.
⋮

2 Harris (1946) is an accessible introduction to this procedure.

(8) * The swim exists.
* The have exists.
* The ate exists.
* The broke exists.
⋮

(9) * The slowly exists.
* The apparently exists.
* The always exists.
* The decidedly exists.
⋮

(10) * The every exists.
* The much exists.
* The no exists.
* The a exists.
⋮

As can be seen, this technique picks out a list of words that match what the grammar school curriculum calls nouns, and segregates them from the others. A similarly discriminating environment can be devised for each category. For each (major) word class, I've given a distinguishing environment in (11).

(11) a. have ___ eaten: Adverb
b. the ___ thing: Adjective
c. dance ___ it: Preposition
d. in ___ orange: Determiner
e. must ___ there: Verb

Understand (11), and (4) as well, as abbreviating the following claim: there is a sentence that is grammatical which contains "X ___ Y," and for which replacing a word of category category into "___" uniquely preserves grammaticality. So, for instance, (11a) should be understood as claiming that all the ways of completing the sentence in (12) involve filling "___" with an adverb.

(12) They have ___ eaten rutabagas.

On this view, morpho-syntactic categories are simply partitions of the vocabulary into equivalence classes. The labels "noun," "verb" and so on are merely convenient names for the resulting subsets of vocabulary items.

There are a few things about the distinguishing environments in (4) and (11) that should be noted. First, they define substitution classes solely on the basis of adjacent items. We might elevate this to a hypothesis.

(13) Morpho-syntactic categories can be defined on the basis of what words they can be adjacent to.

Second, the environments in (11) partition the vocabulary in ways that your language arts curriculum may not have. For instance, the Determiner class picked out by (11d) does not include *much* or *many*. There aren't grammatical sentences that contain *in much thing* or *in many thing* as a substring. One reaction to this would be to allow *much* and *many* to belong to different word classes than *every*, *the*, *a*, and so on. We could admit the two additional word classes, DET_{mass} and DET_{plural}, defined over the environments in (17).

(14) a. in ___ syrup: DET_{mass}
b. in ___ oranges: DET_{plural}

This is a straightforward application of the procedure for defining morpho-syntactic category that Harris's program offers, and it is one direction that syntactic theorists go.

There is another reaction to these data, however, and it is the one I shall follow. It's clear by comparing the environments that define Determiner and Det_{plural} that what distinguishes them is whether the word that follows is plural or singular. The difference between singular and plural is a semantic one, and so we should tie the difference between Determiners and Det_{plural} eventually to a semantic primitive. It is also a semantic difference, although a less familiar one, that distinguishes the Determiner and Det_{mass} categories. Words such as *syrup* refer to entities which do not contain clearly delineated atomic parts, whereas words like *oranges* do. If one recursively divides a group of oranges into its parts, there will come a definitive point — when we are down to the individual oranges — that we will no longer be dividing a group of oranges. A group of oranges is made up of things that one could count. The same is not true of syrup. It is not clear how to find the things that are grouped together to make syrup. Words that refer to entities that can be piled together, or taken apart, in the way that oranges can are called "count nouns," while those that cannot are called "mass nouns." The difference between the Determiner and

Det_{mass} classes is just whether the term that follows them is mass or count. This is a semantic distinction.[3] There is a clearly semantic generalization to be captured in distinguishing these classes of determiners, and we should strive to capture these generalizations in our grammar.

There are generalizations hidden in the environments in (4) and (11) as well, but it is not at all clear that these are semantic generalizations. To see these generalizations, consider the following series of distinguishing environments for the word class 'noun,' each of which is very similar to (4).

(15)
a. the ___ eats
b. some ___ knows
c. a ___ exists
d. few ___ is
e. every ___ ate
f. no ___ exists
g. some ___ has
h. every ___ put
i. a ___ screamed
j. few ___ drove
k. and so on

The generalization in this list is that the words flanking the environment in which nouns are restricted are themselves of a word class; each member of this list fits the schema in (16).

(16) DETERMINER ___ VERB

Each of the environments in (11) can be similarly converted into a generalization that makes reference to morpho-syntactic category.

(17)
a. VERB ___ VERB: ADVERB
b. DETERMINER ___ NOUN: ADJECTIVE
c. VERB ___ NOUN: PREPOSITION

3 But it is still, apparently, one that arbitrarily maps onto lexical items, as Brendan Gillon reminds me. English decides to treat *furniture* as a mass noun, for instance, even though a group of furniture is, arguably, composed of clearly delineated parts. The point here is that there is a way of defining the mass/count distinction. It doesn't need to be rendered in terms of substitution classes. There remains the issue of how to align this semantic definition of the distinction with our, hopefully, evolving folk metaphysics.

d. PREPOSITION ___ NOUN: DETERMINER
e. INFL ___ PREPOSITION: VERB

(NB: The word *must* belongs to a morpho-syntactic category with a small set of members; I've labeled it INFL in (17e). We'll soon encounter the evidence for this category.) At present it is not possible to reduce this generalization to a semantic one. That is, there is no known method of defining morpho-syntactic categories in semantic terms.[4] At present, the best that can be done is to define morpho-syntactic categories in the terms that Zellig Harris gave us: substitution classes. The generalizations underlying (4) and (11) are at present irreducibly morpho-syntactic, then.

Notice that converting (11) to (17) claims that the particular lexical items chosen will not matter. But, as we've just seen, it does matter: whether the noun in (17d) is count or mass or singular or plural will determine which of the Determiner, Det_{mass} and Det_{plural} classes are well-formed in this position. To take the step that defines substitution classes in terms of other substitution classes, then, requires factoring out the semantic information and introducing, as a consequence, certain ungrammatical strings.

One reaction to the differences among Determiner, Det_{mass} and Det_{plural}, then, is to segregate the kinds of information that together determine the distribution of words into a syntactic component and a semantic component. This is the path we shall take. We assign to the semanticist the task of explaining the wholly semantic part of this job: why, for instance, *much* can be left-adjacent to a mass noun but not a count noun. In general, it is not trivial to know when it is the semantics or the syntax that is responsible for coöccurence restrictions like those in (4) and (11), and the line is constantly being questioned. Harris, it seems, believed that virtually none of it was semantic, whereas present-day categorial grammarians push in the direction of removing an independent syntactic contribution. I'll chart a course that is somewhere in the middle.

Morpho-syntactic categories, then, are defined syntactically. They are subsets of the vocabulary that can be substituted for each other in particular positions within a grammatical sentence preserving grammaticality. Moreover, the particular positions can be characterized in terms of adjacent morpho-syntactic categories. The first step in characterizing the grammaticality judgments of some speaker is recognizing that the vocabulary of that speaker is partitioned in this way.

4 For a recent attempt to define some of the major categories in terms that verge on semantic, see Baker (2003).

2.2 Phrases

Certain strings of categories also have a distribution within sentences that can be defined in terms of adjacent items. For example, the string D(eterminer)+Adj(ective)+N(oun) can appear immediately after a preposition and immediately preceding the *'s* which marks the "possessive."

(18) ___ *'s* & P ___ : D+Adj+N
  a. I talked to **the happy woman**.
  b. **the happy woman**'s friend

This string can also be "coördinated" with another identical string of categories. Coördination involves the use of words called "conjuncts," words such as *and*, *or*, *nor*, etc. Thus, we find examples like (19) but not (20).

(19) the happy woman and an unhappy man

(20) a. * the angry and an unhappy man
  b. * the and an unhappy man

Finally, with respect to all these distributional tests, the strings D+N+P(reposition)+N, N+P+N, Adj+N, N, and (infinitely) many others also pass. We need some way of describing the fact that these strings are "the same," and different from, say, P+N which has a distinct distributional pattern. That is, this family of strings is a substitution class in the same sense that morphosyntactic categories are.

Families of strings like this are called "phrases," and we can write a Phrase Structure Rule to describe which strings belong to such a family. In the case at hand, this rule might look like (21).

(21) αP → (D) (Adj) N

Understand material enclosed within "( )" to be optional; (21) therefore generates the set of strings: D+Adj+N, D+N, Adj+N and N.

This leaves out the strings D+N+P+N and N+P+N. But these strings involve another phrase, made up of the string P+N. This string, along with any string that conforms to the template P or P+αP or P+P or P+P+αP has the defining distribution in (22).

(22) A ___ & P ___
  a. I stood **around**.
  b. I knew the man **by Mary**.

c. I remain disliked **by Mary**.

d. I stood next **to Mary**.

Like αPs, βPs may be coördinated with other βPs, but not with other αPs, as the following examples illustrate.

(23) a. Under the bed and behind the sofa are usually good places to find money in my house.

b. * The dining room table and behind the sofa are usually good places to find money in house.

Hence, just as with αPs, this family of strings constitutes a substitution class.

Putting these observations together, we come up with the Phrase Structure rules in (24).

(24) a. $\alpha\text{P} \rightarrow \text{(D) (Adj) N (}\beta\text{P)}$

b. $\beta\text{P} \rightarrow \text{P (}\alpha\text{P)}$

c. $\beta\text{P} \rightarrow \text{P (}\beta\text{P)}$

It is customary to collapse the two rules in (24b,c) to (25).

(25) $$\beta\text{P} \rightarrow \text{P} \left\{ \begin{array}{c} (\alpha\text{P}) \\ (\beta\text{P}) \end{array} \right\}$$

Material enclosed in "{ }" offers a set of exclusive choices: exactly one of the members of the enclosed list must occur. In (25) this means that either αP, βP or, because these are both optional, nothing may occur after P to form a PP.

Note that (24a) and (25) together have the property of being recursive. This is an important aspect of phrase structure rules for it is the primary means by which we describe the indefinite length, and cardinality, of sentences. These two phrase structure rules are able to characterize infinitely many and infinitely long strings of words. This is a correct result, for we are, as far as our linguistic abilities go, capable of forming grammaticality judgments about an infinity of sentences and about sentences of infinite length. We have other properties that prevent us from doing this, of course. Our memories and attention are too ephemeral for even very long sentences; and even if we were able to overcome these cognitive limitations, our death will bring an eventual end to any sentence, or series of sentences, that we are evaluating. But there is no reason to think that this limitation is a linguistic one. We should let our model of grammaticality judgments characterize an infinity of sentences, as well as permit sentences of infinite length, and let the actual limits on the lengths and

(29) a. γP → V (αP) (βP)

b. $\beta P \rightarrow P \left\{ \begin{array}{c} (\alpha P) \\ (\beta P) \end{array} \right\}$

c. αP → (D) (Adj) N (βP)

numbers of sentences that we evaluate be determined by other factors. The recursiveness of phrase structure rules is a step in that direction.

Still another phrase structure rule is required to account for the fact that the family of strings that include V, V+αP, V+βP, V+αP+βP, and an infinite set of other such strings is a substitution class. The environment that defines them is (26).

(26) Infl ___

a. I should **eat rutabagas**.

b. I will **talk to Mary**.

c. I will **tell Mary about rutabagas**.

"Infl" is a morpho-syntactic category that includes *should*, *will*, *must*, *would*, *can*, *could* and a few other words. It's an abbreviation for "inflection"; the words that belong to it are tied to the inflectional classes that verbs belong to, as we shall see. Like αPs and βPs, coördination treats members of this family as equivalent and distinct from αPs and βPs.

(27) a. Mary walked and talked.

b. Mary visited Paul and kissed Barry.

c. Mary talked to Paul and met with Barry.

These facts call for a Phrase Structure rule like the following:

(28) γP → V (αP) (βP)

We have now arrived at the three rules in (29) below. There is a common property to all these rules. In each case, all of the constituents are optional, except one. Thus, a verb is the only necessary member of a γP, a noun the only requisite member of an αP and a preposition is all that's required to make a βP. This is just another way of observing that the environments that define these phrases are also environments in which a word class is defined. Further, the converse also turns out to be true: whenever there is a preposition, there is a βP, wherever a noun is found, there is an NP, as so on. More precisely, the

environments that define a phrase will always include an environment that defines some category. Thus, nouns and αP, prepositions and βP, verbs and γP are in one-to-one correspondence. This is a very pervasive property of Phrase Structure rules. Phrase Structure rules vary to a considerable degree across languages, but this property of them seems to always hold. We'll confront two apparent counterexamples from English shortly, but these are probably only apparent counterexamples. So far as I am aware, there is no clear counterexample to this generalization. This property of Phrase Structure rules is known as *endocentricity*. The word that must be a member of the phrase is its *head*. Finally, it is common practice to name the phrases after their heads, so we'll rename αP, NP, βP PP and γP VP. Thus, we now have the rules in (30).

(30) a. NP → (Det) (Adj) N (PP)

b. $\text{PP} \rightarrow \text{P} \left\{ \begin{array}{c} \text{(NP)} \\ \text{(PP)} \end{array} \right\}$

c. VP → V (NP) (PP)

In addition to these three Phrase Structure Rules, we'll need quite a few others. Indeed, the principle of endocentricity leads us to expect that for every category, there will be a Phrase Structure rule that builds a phrase headed by that category. For example, corresponding to the category Adjective, there is a rule that builds adjective phrases; (31) is a good first approximation.

(31) AP → A (PP)

The presence of PPs within Adjective phrases is supported by the existence of strings like:

(32) a. She is **interested in syntax**.
She is **interested.**

b. He seems **happy with linguistics**.
He seems **happy**.

The coördination test also treats A and A+PP strings as being the same, as (33) indicates.

(33) a. She is happy and interested in syntax.

b. He seems bored but happy with linguistics.

We'll also need a Phrase Structure rule that tells us how these various phrases are put together to form a sentence. (34) looks roughly right.

(34) S → NP Infl VP

The morpho-syntactic category that sentences are in a one-to-one relation with is Infl,[5] and so in keeping with the convention of naming phrases after their heads, we should change (34) to (35).

(35) IP → NP Infl VP

With this rule we have finally come to the task of characterizing the grammaticality judgments of English speakers. For any speaker of English whose vocabulary has been partitioned into noun, verb, preposition, adjective, determiner and Infl, (35), with the rules in (30), characterizes those strings of words that will be judged grammatical.

This is just a first step, of course. We have hundreds of pages left. In fact, it's possible to see something wrong with (35) right away. It says that no sentence can fail to have an Infl between NP and VP, but if Infl are just words such as *can*, *could*, *will*, and so on this is obviously wrong. There are grammatical sentences aplenty that fail to have these words in them. (36) is an example.

(36) Jerry walked.

Where is the Infl between *Jerry* and *walked* in this sentence?

If we look hard, we find that sentences are, in fact, in a one-to-one correlation with a category, but that category includes not just words, but bound morphemes as well. Consider the sentences in (37).

(37)
a. Jerry leave**s**.
b. Sally lef**t**.
c. Sam ha**s** left.
d. Sarah ha**d** left.
e. Martha **should** leave.
f. George **might** have left.
g. Laura desires [Sal **to** leave].
h. Larry remembers [Jim leav**ing**].

5 In this case, however, unlike what we found for the VP, NP and AP rules, Infl is not the only obligatory member of a sentence. It is presently controversial whether sentences are the only phrases that have this property. We will see in later lectures that there are syntactic reasons for the obligatory presence of NP, and likely semantic ones for the obligatory presence of VP.

The boldfaced terms have similar distributions: they are found either immediately preceding the verb (if they are free) or affixed onto the following verb (if they are bound). Every sentence has one of these, and so these terms meet the criteria of being the head of a sentence. To explain how it is that those Infls which are bound morphemes materialize affixed onto the following verb, we will have to invoke a process that goes beyond phrase structure rules. Let us put off doing this.

As we gather more detail about the shapes of grammatical English sentences, we will need to make quite a number of additions to these rules. In fact, to be exhaustive about this proves to be a task beyond what we can manage here; we should consider this an open-ended process. Nonetheless, I want to gather a little more detail than we now have.

I'll begin by adding a couple of phrases to our inventory. One of these is a sort of "sentence" found in examples like (38).

(38) a. Mary said **that John likes chocolate**.
b. Mary recalled the rumor **that John likes chocolate**.
c. **That John likes chocolate** bothers Mary.
d. Jerry is angry **that John likes chocolate**.

Note that the strings following the word *that* meet the conditions imposed by the rule that builds IPs. The word *that* is called a "complementizer" and it is the head of the phrase found in these sentences. This phrase, or clause as sentence-like phrases are often called, is a "Complementizer Phrase" (CP). CPs conform to the requirements of the following Phrase Structure rule.[6]

(39) CP → C IP

Other complementizers are *if* and *whether*, as found in the following examples.

(40) a. I wonder if Mary likes chocolate.
b. I asked whether Mary likes chocolate.

Having introduced this constituent, we will now need to revise our previous Phrase Structure rules to include the positions where they may lie. This yields the following battery of rules.

6 The conclusion that sentences are headed by Infl, and that the subordinate sentences that include *that* are headed by complementizers, is reached in Chomsky (1986a). In the generative literature that precedes this work, these two types of sentences were given various treatments. Bresnan (1972) suggested that *that*-clauses were of the same category as sentences; she called them $\overline{\text{S}}$s. Jackendoff (1977) argued that sentences are larger VPs, headed by a verb. The IP/CP model is now standard, though the issue is periodically revisited.

(41) a. $IP \rightarrow \left\{ \begin{array}{c} NP \\ CP \end{array} \right\} I\ VP$

b. NP → (D) (AP) N (PP) (CP)

c. VP → V (NP) (PP) (CP)

d. AP → A (PP) (CP)

Note the option of having a CP in place of an NP at the beginning of a sentence. Note too that I've brought the NP rule into conformity with the principle of endocentricity. Our earlier rule (i.e., NP → (D) (Adj) N (PP)) permitted an adjective without an adjective phrase. I've replaced "(Adj)" with "(AP)" to correct for this. We'll see the empirical support for that change shortly.

The second phrase we'll need are ones headed by adverbs. Adverbs are a word class that is sometimes defined on the position left adjacent to adjectives, as in (42).

(42) a. A **very** happy child

b. The **extremely** large boat

They can also be found left adjacent to verbs, as in (43).

(43) a. I have **deliberately** misled.

b. I have **noticeably** erred.

Interestingly, to a large extent, the set of adverbs that can be immediately preceding a verb is a proper subset of those that can immediately precede an adjective.

(44) a. A **deliberately** angry child

b. The **noticeably** large boat

(45) a. * I have **very** misled.

b. * I have **extremely** erred.

Here is one of those places where we must judge whether we have two morpho-syntactic categories or we have a semantic distinction. To judge from these few examples, it seems that the adverbs which indicate an "extent" or "measure" are fine in the pre-adjectival position but not in the preverbal position. This description of the contrast makes reference to the meanings of the words involved, and so it could reflect a semantic fact. On the other hand, it may be that we **are** looking at a contrast due to morpho-syntactic category, but one that correlates with (or perhaps is defined by) these semantic factors. I don't

know how to decide between these alternatives. I will gamble that this is a category distinction like the others we are examining. Let's call the category of words that I've characterized as denoting an extent, "degree," and reserve the term "adverb" for the others.

The class of Degree words can also show up in combination with adverbs, as illustrated by (46).

(46) a. Mary **very** quickly walked through the full classroom.
b. Mary **extremely** loudly declared her interest in Dutch phonemics.

Indeed, wherever an Adverb can be found, so also can a Deg+Adverb string. Our phrase structure rule for adverb phrases will be (47), therefore.

(47) AdvP → (Deg) Adv

If we admit the category Degree, then the principle of endocentricity tells us that there must be a phrase that has the same distribution. It is very difficult to identify a (non-trivial) string that has the same distribution as Degree words. Those strings are most easily seen in so-called comparative constructions, a complex phenomenon that we will hopefully manage to skirt in these lectures. As a stop-gap, let us do with the questionable rule in (48).

(48) DegP → Deg

We should modify (47) accordingly.

(49) AdvP → (DegP) Adv

The existence of degree phrases and adverb phrases now requires that we modify some of our other rules so that they can be positioned within sentences correctly. Adjective phrases will have to be changed to allow for degree phrases within them, as in (50).

(50) AP → (DegP) A (PP) (CP)

And refinements to the rule that characterizes verb phrases will be necessary in order to position adverb phrases within them.

(51) VP → (AdvP) V (NP) (PP) (CP)

There are some other changes to the VP rule that are necessary. Note, for instance, that VPs may occur immediately following a verb, as in (52).

(52) a. Mary has walked.
b. Mary has talked to John.

c. Mary has visited Gary.

Interestingly, if the verb heading a VP is followed by another VP, nothing else may follow the head verb. For instance, *Mary has on the platform walked* is ungrammatical. We need, therefore, to modify the VP Phrase Structure rule in such a way that the head verb is followed by a VP, or by the expansion previously arrived at, but no combination thereof. This can be done with the aid of curly brackets in the following way:

(53) $$\text{VP} \rightarrow \text{(AdvP) V} \left\{ \begin{array}{c} \text{(NP)} \quad \text{(PP)} \quad \text{(CP)} \\ \text{VP} \end{array} \right\}$$

Further, it is possible to find APs embedded within VPs; (54) provides some examples.

(54) a. Sally remains angry at Jim.
b. Frank is happy with himself.

When APs follow verbs, they may be preceded by, at most, a PP, as in (55).

(55) Jerry seems [$_{\text{PP}}$ to Bill] [$_{\text{AP}}$ happy with his rutabagas].

So we change the rule that characterizes VPs to:

(56) $$\text{VP} \rightarrow \text{(AdvP) V} \left\{ \begin{array}{c} \text{(NP)} \quad \text{(PP)} \quad \text{(CP)} \\ \text{VP} \\ \text{(PP) AP} \end{array} \right\}$$

Finally, consider that part of the NP rule that introduces determiners. Determiners include words like *the*, *a*, *that* (not to be confused with the complementizer *that*), *every*, *some*, *all*, etc. Interestingly, it's very rare that we find determiners combining with other words to form a phrase that combines with a following noun. A couple of examples which might constitute cases of this sort are given in (57).

(57) a. all but three dogs
b. more than most people

I don't know precisely what the Phrase Structure rule is that determines which strings may stand in this position. Nonetheless, one common approach to these cases is to imagine that determiners head their own anemic phrases, which are then positioned within NPs. We will revisit this idea, but for now let's imagine that determiner phrases are made up of nothing but determiners.

(58) DP → Det

We'll therefore need to update the phrase structure rule that forms NPs. But before we do this, let's consider strings like those in (59).

(59) a. Mary's book
b. the man's toy
c. the man on the table's nose

These examples involve a possessive or genitive phrase. Note that this phrase is an NP with the morpheme *'s* appended to the end. Further, note that this genitive phrase never co-occurs with a DP, as (60) illustrates.

(60) a. * the Mary's book
b. * the the man's toy
c. * a the man on the moon's nose

One very typical explanation for this is to understand determiners and possessives as competing for the same position. In this situation, that can be done by rigging the NP phrase structure rule in such a way that it either produces a DP or a genitive phrase in the same position. This is done with the curly braces abbreviation in (61).

(61) $\text{NP} \rightarrow \left\{ \begin{array}{c} \text{(DP)} \\ \text{(NP's)} \end{array} \right\} \text{(AP) N (PP) (CP)}$

One final Phrase Structure rule is required by the sorts of examples we've so far reviewed. This is the Phrase Structure rule that generates coördinated phrases. This can be done with the following.

(62) $\alpha \rightarrow \alpha$ Conj α

This rule says that a phrase of any category can be made up of two other such phrases with a conjunct stuck between them. Conjuncts, recall, are *and*, *or* and *but*.

Summarizing, we've now introduced the following battery of Phrase Structure rules:

(63) a. $\text{IP} \rightarrow \left\{ \begin{array}{c} \text{NP} \\ \text{CP} \end{array} \right\} \text{I VP}$

b. $\text{NP} \rightarrow \left\{ \begin{array}{c} \text{(DP)} \\ \text{(NP's)} \end{array} \right\} \text{(AP) N (PP) (CP)}$

c. $VP \rightarrow (AdvP)\ V \left\{ \begin{array}{c} (NP) \quad (PP) \quad (CP) \\ VP \\ (PP)\ AP \end{array} \right\}$

d. DP → Det

e. DegP → Deg

f. AdvP → (DegP) Adv

g. AP → (DegP) A (PP) (CP)

h. CP → C IP

i. $PP \rightarrow P \left\{ \begin{array}{c} (NP) \\ (PP) \end{array} \right\}$

j. $\alpha \rightarrow \alpha$ Conj α

An interesting property of the phrases defined in (63), a property which Harris discussed, is that they all distinguish the phrases they are defining from the word class which that phrase matches. One might wonder why there is this distinction. We have decided that the way to define category is as classes of words that can substitute into a certain collection of positions. We've also used this very definition for phrases. And the principle of endocentricity tells us that there is a one-to-one correspondence between category and kind of phrase. What we end up with, then, is a situation where the positions that define a phrase are always a proper subset of the positions that define a category. We might wonder why the positions defining a phrase are always a proper subset of those that define a word class; why aren't they the very same positions? Why, for instance, isn't the rule for noun phrases something like (65) rather than (65)?

(64) NP → (DP) (AP) N (PP)

(65) N → (DP) (AP) N (PP)

Harris argues that phrases need to be distinguished from word class, and that this is a general property of phrase structure rules. He points out, for example, that while singular nouns are in the same substitution class as are plural ones, a plural noun cannot substitute for a singular one when it combines with the plural morpheme. He assumes that phrase structure rules control inflectional morphology, and therefore, that in addition to the rules we have discovered, there is also a rule such as (66) that produces plural nouns.

(66) NP → N s

(We will adopt something similar in a few chapters.) His point, then, is that this rule should not be (67) because that would wrongly produce doubly pluralized nouns such as (68).

(67) N → N s

(68) kitses [kɪtsɨz]

This is the reason, then, why our phrase structure rules look like (64) and not (65). Or, to put it somewhat differently, we do not want these rules to be recursive with respect to their head. The phrases we've encountered so far all have this property.

But interestingly, it turns out that not all phrases do. Some phrases are headed by other phrases. And these phrases, it turns out, are identical. We turn to these cases next.

2.3 $\overline{X}$ phrases

There are substitution classes that pick out strings which are recursive on themselves. That is: these phrases are headed by themselves. These phrases are found inside those we've identified so far. For example, in the position marked by "___" in (69), we find the family of strings in (70). Some examples are in (71).

(69) Det ___ V

(70) {N, AP N, N AP, N PP, AP N PP, N PP AP, AP AP N, N PP PP, AP AP N PP, AP AP N PP PP, . . . }

(71) the woman left.
the happy woman left.
the woman unhappy with the lecture left.
the happy woman with a hat left.
the woman with a hat unhappy with the lecture left.
⋮

And coördination also reveals that this set of strings forms a family:

(72) The woman and happy man left.
The happy woman and man with a hat left.
⋮

Now this family of strings does not appear to be the family we have called NP. There are two, related, reasons for this. First: there are grammatical strings from the second family which cannot be substituted for instances of the first family, as (73) indicates.

(73) a. The woman left.
b. * Woman left.

Second: a close inspection of the set that the second family is made up of indicates that it does not share Harris's property. This family is recursive with respect to itself. Unlike NPs, which can have only one instance of a DP inside them, the phrase we've discovered here can contain any number of strings of the same kind as itself. So, we set up something like (74).

(74) a. $\text{NP} \rightarrow \left\{ \begin{array}{c} \text{(NP's)} \\ \text{(DetP)} \end{array} \right\} \overline{\text{N}}$
b. $\overline{\text{N}} \rightarrow \text{AP}\ \overline{\text{N}}$
c. $\overline{\text{N}} \rightarrow \overline{\text{N}}\ \text{AP}$
d. $\overline{\text{N}} \rightarrow \overline{\text{N}}\ \text{PP}$
e. $\overline{\text{N}} \rightarrow \text{N}$

Note how these rules encode the "optionality" of AP and PP differently than the optionality of DP. And note, further, that they are all endocentric on N. They also leave out the position of CP; this is because fitting CPs into this structure poses a problem. We will return to it in just a moment.

We find the existence of very similar subphrases within VPs as well. Consider, for instance, the environment in (75), which permits the family of strings in (76), as (77) exemplifies.

(75) NP ___ CP

(76) {V, AdvP V, V AdvP, V PP, AdvP V PP, V PP AdvP, V AdvP PP, AdvP V PP PP, AdvP AdvP V, AdvP AdvP V PP, V PP PP, …}

(77) Sally said that Jerry left.
Sally quickly said that Jerry left.
Sally quickly said to Peter that Jerry left.
Sally said to Peter quickly that Jerry left.
Sally said quickly to Peter that Jerry left.
Sally carefully said to Peter on Tuesday that Jerry left.
⋮

And, as before, coördination recognizes this family.

(78) Sally **shouted** and **whispered** that Jerry left.
Sally **loudly shouted** and **whispered** that Jerry left.
Sally **shouted to Peter** and **quietly whispered** that Jerry left.
⋮

Again, this subphrase is recursive and headed. So we have something like:

(79)
a. VP → $\overline{\text{V}}$
b. $\overline{\text{V}}$ → AdvP $\overline{\text{V}}$
c. $\overline{\text{V}}$ → $\overline{\text{V}}$ AdvP
d. $\overline{\text{V}}$ → $\overline{\text{V}}$ PP
e. $\overline{\text{V}}$ → V

These rules leave out the expansions of VP which introduce NPs, CPs, APs, and VPs. Moreover, the first of these rules says that VPs are $\overline{\text{V}}$s and nothing more, which raises the obvious question why we posit $\overline{\text{V}}$s here at all. We would get the same result by dispensing with the first of these rules, and replacing $\overline{\text{V}}$ with VP throughout the remainder. We will soon see, however, that in certain situations there is a term that can show up which appears to be dominated by VP but not $\overline{\text{V}}$. I'll keep these rules in anticipation of that situation.

A similar situation arises in Adjective Phrases too. If we examine the environment in (80) we discover that it characterizes the set of strings in (81).

(80) V ___ CP

(81) {A, Deg A, Deg Deg A, A PP, Deg A PP, Deg A PP PP, ...}

(82) Sean is **happy** that syntax is cool.
Sean was **happy on Tuesday** that syntax is cool.
Sean was **very happy on Tuesday in this class** that syntax is cool.
⋮

As before, this family is recursive and headed. And, as before, it is visible to coördination as well.

(83) A child **happy with her guardian** and **well-rested** is unlikely to cause trouble.
A child **happy with her guardian** and **completely well-rested** is unlikely to cause trouble.
A child **thoroughly unhappy in a zoo** and **angry at her guardian** is

likely to cause trouble.

⋮

We need to revise the AP rule to something like:

(84) a. AP → $\overline{\text{A}}$
 b. $\overline{\text{A}}$ → Deg $\overline{\text{A}}$
 c. $\overline{\text{A}}$ → $\overline{\text{A}}$ PP
 d. $\overline{\text{A}}$ → A

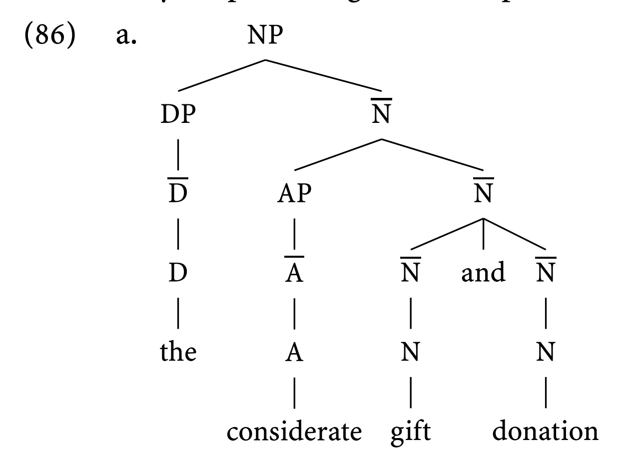
Note that I have left out CP, as in the other rules; and, as with the VP rule, these rules characterize AP as consisting of just an $\overline{\text{A}}$ and nothing else. Both matters we'll take up shortly.

There is a feature of this method of representing these subfamilies that I would like to draw attention to now. It allows for two separate parses of examples such as (85).

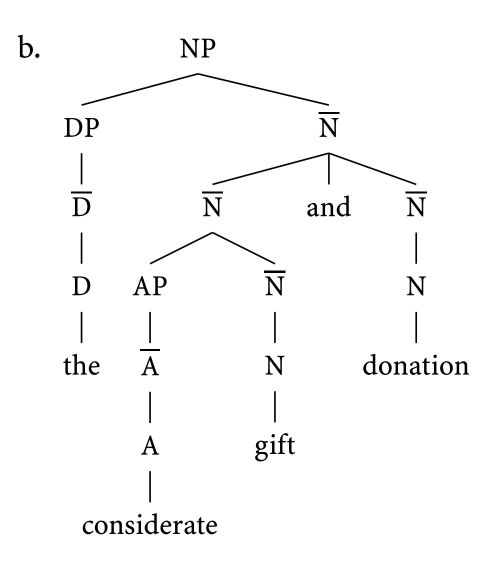
(85) the considerate gift and donation

It is possible to produce this string either by grouping *considerate* and *gift* into one $\overline{\text{N}}$ and conjoining that with an $\overline{\text{N}}$ consisting of just *donation*, or it is possible to conjoin *gift* and *donation* into one $\overline{\text{N}}$ and then group that phrase with *considerate* into an $\overline{\text{N}}$. It is easy to represent these two parses by way of "phrase marker trees," which graphically elucidate the constituent structure of strings. The two ways of producing (85) are represented by the trees in (86) below.

(86) a.



b.



We might note that there are two meanings attached to this string as well, having to do with how the meaning of *considerate* is combined with the meanings of the rest of the parts. A loose paraphrase of these two meanings might be as given in (87).

(87) a. the things which are considerate and which are, first, a gift and, second, a donation

b. the things which are, first, a considerate gift and, second, a donation

There is some reason for thinking that these two syntactic representations map onto those two interpretations. For one thing, the number of meanings and the number of parses matches. For instance, if we add one more adjective to the left of the coördinated nouns, as in (88), our rules allow for a total of three parses (shown in (90) on the following page) and there are three meanings as well (as indicated in (89)).

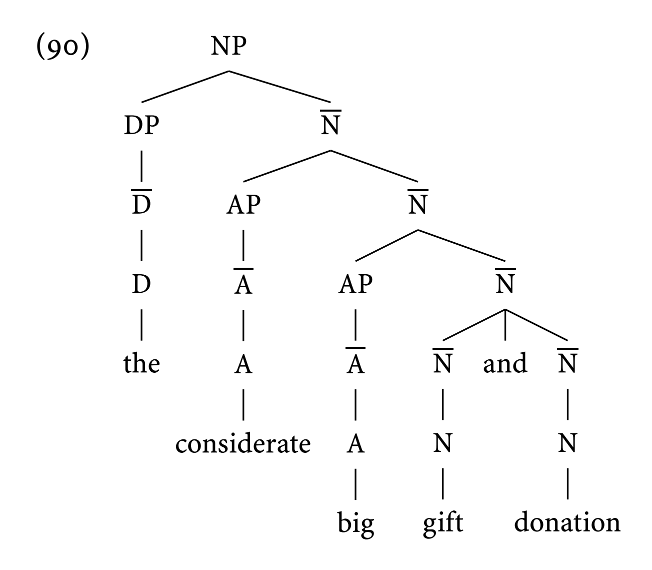
(88) the considerate big gift and donation

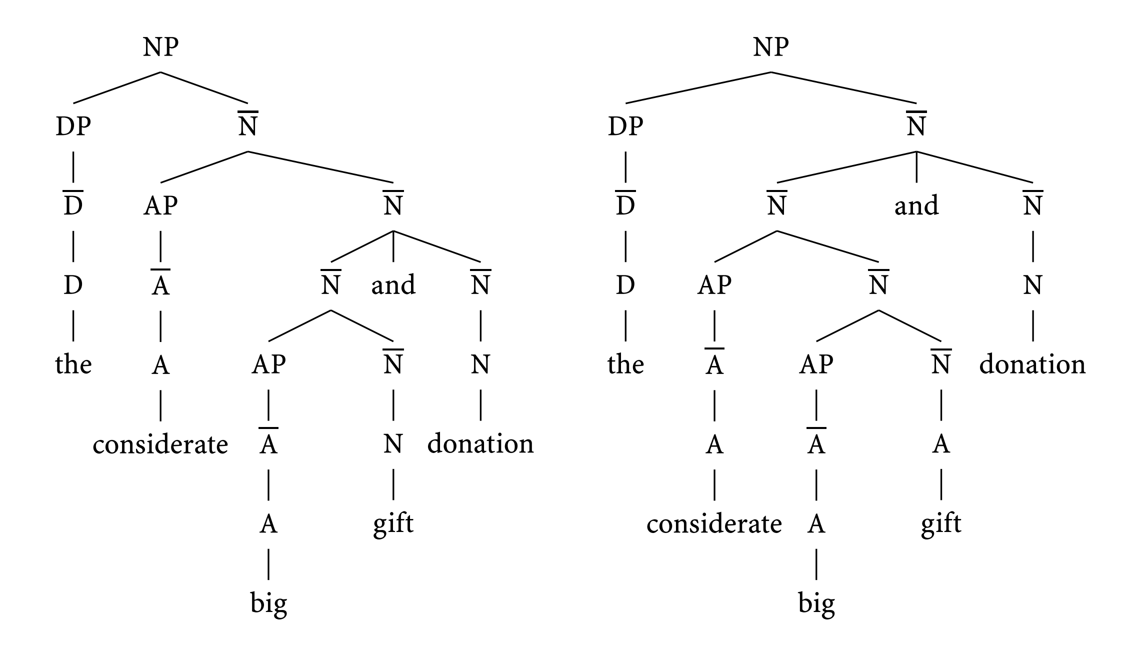
(89) a. the things which are considerate and big and are also a gift and a donation.

b. the things which are considerate and are also a big gift and a donation

c. the things which are a considerate big gift and a donation

(90)


NP
DP
D̄
D
the
N̄
AP
Ā
A
considerate
N̄
AP
Ā
A
big
N̄
N̄
and
N̄
N
gift
N
donation


NP
DP
D̄
D
the
N̄
AP
Ā
A
considerate
N̄
N̄
and
N̄
AP
Ā
A
big
N̄
N
gift
N
donation
NP
DP
D̄
D
the
N̄
N̄
and
N̄
AP
Ā
A
considerate
N̄
AP
Ā
A
big
N̄
A
gift
N
donation

Furthermore, the meanings vary in a predictable way with the linear order that these terms are arranged in. Thus, for instance, putting the second adjective to the left of the coördinated nouns creates the three meanings listed in (89), whereas putting the second adjective to the left of just the rightmost noun, as in (91), produces just two readings: they are (92).

(91) the considerate gift and big donation

(92) a. the things which are considerate and both a gift and a big donation
b. the things which are a considerate gift and a big donation

This is predictable in the sense that our characterization of these strings would deliver just the two parses for (91) shown in (93) on the next page.

This correspondence should give us some courage that we are on the right track in characterizing the infinite strings under discussion in terms of recursive phrases. It provides a set of structures that are in correspondence with what look like a parallel set of meanings. Our next step should be to flesh out this correspondence, but we have some work still to do in characterizing these basic facts about grammaticality judgments. So let's return to that task.

The strings belonging to Adverb Phrases are so simple that it is difficult to know whether they contain the substructure we've found in the other phrases. Nonetheless, they do have a recursive part and this might be construed, on analogy with these other cases, as evidence for substructure:

(94) Sally carefully spoke.
Sally very carefully spoke.
Sally very, very carefully spoke.
⋮

The coördination phenomenon also seems to suggest subphrases, at least if our decision about the meaning-form mapping made above is correct.

(95) Sally spoke [almost [very rapidly] and [quite softly]].

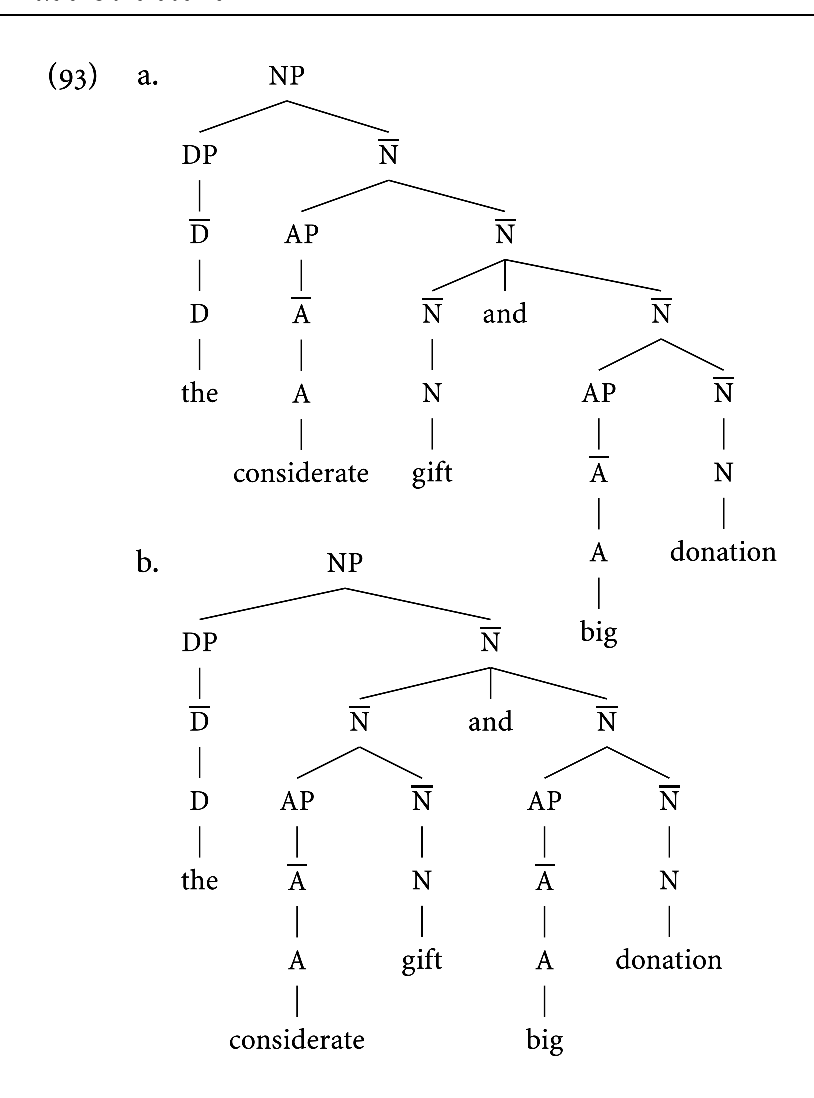
So, let's convert the AdvP rule to (96).

(96) a. AdvP → $\overline{\text{Adv}}$
b. $\overline{\text{Adv}}$ → DegP $\overline{\text{Adv}}$
c. $\overline{\text{Adv}}$ → Adv

Like the AP and VP rules, this battery of rules equates AdvP with $\overline{\text{Adv}}$ and so makes mysterious why they are called different things.

The rule building sentences, IPs, is similarly meager. But it too shows some signs of the subfamilies which we have discovered in NPs, APs and VPs. This is indicated by coördination in examples such as (97).

(93) a.



(97) Jerry [can speak loudly] but [can't speak clearly].

And, when we add to our observations that adverbs can fall to the left of Infl, we discover the recursive flag of these intermediate families:[7]

(98) Jerry evidently won't speak.
Jerry evidently deliberately won't speak.

7 But haven't we already characterized strings like the third and the fifth in (98) as adverbs introduced by a recursive $\overline{\text{V}}$ rule? Do we really need to also let sentence-final adverbs be introduced by a recursive $\overline{\text{I}}$ rule? The answer typically given is: yes. But you might want to decide for yourself what the answer to this should be.

Jerry evidently won't speak deliberately.
Jerry evidently occasionally deliberately won't speak.
Jerry evidently won't speak occasionally deliberately.

(These are all somewhat strained, I grant you, but I think still grammatical.) This calls for a change along the lines in (99).

(99) a. $IP \rightarrow \left\{ \begin{array}{c} NP \\ CP \end{array} \right\} \bar{I}$

b. $\bar{I} \rightarrow AdvP\ \bar{I}$

c. $\bar{I} \rightarrow \bar{I}\ AdvP$

d. $\bar{I} \rightarrow I\ VP$

Note how in this battery of rules, unlike the others we've formulated, the $\overline{X}$ rule that terminates the recursion has more than just the "head" of the phrase in it. In this case it also introduces the VP. This is required because VPs are not recursively introduced, and the method we have adopted of representing recursion in these phrases is built into the structure of the substitution classes.

Actually something similar is true for the rules that build APs, NPs and VPs as well. In the case of VPs, the NP and CP parts of their family are not recursively introduced. So we should change the terminal expansion to:

(100) $\overline{V} \rightarrow V\ (NP)\ (CP)$

And similarly, the CP parts of the AP and NP families are not recursively introduced, so the terminal expansions of these families should be changed to:

(101) $\overline{A} \rightarrow A\ (CP)$
$\overline{N} \rightarrow N\ (CP)$

So this corrects the omission of CP and NP in our original formulation of these rules, though, as foreshadowed above, this will produce a difficulty.

To see this difficulty, consider how our structural method of stopping the recursion relates the terms that are within some phrase. We expect that those terms which are introduced in the terminal expansion "$\overline{X} \rightarrow X \ldots$" (that is, the non-recursively introduced terms) will form the most inclusive substitution class of the phrase involved. There are some kinds of phenomena which suggest that this expectation is fulfilled. There are processes, for example, in which a rather surprisingly short string can substitute for one or another of the families we have discovered. This happens under conditions of anaphora.[8]

8 "Anaphora" refers to processes in which a phrase in one position refers, in some fashion or other, to the same things that another phrase, in a different position, refers to. For instance, in

For example, the term *one* prosodically looks like a word, but semantically derives its meaning by being anaphoric to an $\overline{\text{N}}$.

(102) a. I will examine the blue book about language if you will examine the brown one.
one = "book about language"

b. I will examine the big blue book about language if you will examine the small one.
one = "blue book about language"

c. I will examine the long book about language if you will examine the one about Quarks.
one = "long book"

The reason we think that *one* not only semantically is an $\overline{\text{N}}$, but is also syntactically an $\overline{\text{N}}$, is because of the contrast in (103), a contrast which also supports our treatment of the non-recursive parts of NP.

(103) a. I will examine the long proof that language exists if you will examine the short one.
one = "proof that language exists"

b. * I will examine the long proof that language exists if you will examine the one that it doesn't.
one = "long proof"

What this contrast indicates is that *one* must "stand in" for the noun and the CP that follows, and cannot stand in for the noun by itself. This is explained if *one* can stand in for an $\overline{\text{N}}$, because there is no $\overline{\text{N}}$ under the current rule set that fails to contain these both. It isn't, incidentally, that there is some semantic constraint on *one* that prevents it from standing in for something that has the meaning of a single noun, because that is possible in cases such as (104).

(104) I will examine the book on the shelf if you will examine the one on the table.

The difference between (104) and (103b) is just whether the material that combines with *one* is allowed to be a sister to an $\overline{\text{N}}$ or not: PPs are (look at (102c)), and CPs aren't.[9]

the sentence: *Mary knows that she is smart*, it is possible for *she* to refer to the same individual that *Mary* refers to. In such a case, we say that *she* is **anaphoric** to, or with, *Mary*.

9 We will see, in just a moment, that this does not turn out to be a distinction that hinges on the PP/CP difference, however — so be forewarned. In particular, it will emerge that the semantic

Similarly, the $\overline{\text{V}}$ family can be anaphorically connected to other $\overline{\text{V}}$s, but in this case the phonological manifestation of the anaphor is silence, which will be designated with "Δ" in what follows.

(105) a. Although Sally shouldn't Δ, Jerry must leave town.
Δ = "leave"
b. Although Sally can carelessly Δ, Jerry must carefully read *Aspects*.
Δ = "read *Aspects*"
c. Because Jerry frantically read *Aspects* after dinner, Sally did Δ just before class.
Δ = "frantically read *Aspects*"

This process of anaphora — called "VP Ellipsis," though it might be more accurate to call it "$\overline{\text{V}}$ Ellipsis" — reveals that the non-recursive parts of the VP family are trapped within the smallest subfamily.

(106) a. * Although Sally shouldn't Δ Chicago, Jerry must leave New York.
Δ = "leave"
b. * Although Sally didn't Δ that she was tired, Jerry said that he would sleep.
Δ = "say"

These processes also indicate that there are at least some PPs that must be part of the terminal expansions of $\overline{\text{N}}$ and $\overline{\text{V}}$.

(107) a. ?? I will listen to this long examination of quarks, if you will listen to the *one* of syntax.
one = "long examination"
b. * Although Sally didn't Δ about George, Jerry will carelessly talk about Sal.
Δ = "carelessly talk"

So we should change these rules to:

(108) $\overline{\text{N}} \rightarrow$ N (PP) (CP)
$\overline{\text{V}} \rightarrow$ V (NP) (PP) (CP)

This way of distinguishing the recursive and non-recursive parts also predicts that the non-recursive parts will always come between the head of their phrase and the recursive parts. This seems true sometimes, as in (109) and (110).

function of the PP or CP determines how it behaves with regard to this test, and not the mere fact that the phrase is a PP or CP.

(109) a. Jill ate it at noon.
b. * Jill ate noon it.

(110) a. Jill ate spätzle at noon.
b. * Jill ate at noon spätzle.

But for other cases it seems uncertain, or downright wrong, as in (111) and (112).

(111) a. Jill ate the rotting kumquats.
b. Jill ate at noon the rotting kumquats.

(112) a. ?? Jill said [that you shouldn't eat kumquats] at noon.
b. Jill said at noon [that you shouldn't eat kumquats].

This then, is the difficulty in trying to place CPs, and certain NPs, within VP (and other phrases too, as we'll see). Let's set this problem aside momentarily. It will be the focus of a lot of our work in the chapters that follow.

There is a similarity to the organization of the family of substitution classes that make up NP, VP, AP, AdvP, and IP. The other phrases: PP, CP, DegP and DP are too anemic for us to see that structure, so we don't know, empirically, whether or not they have it.[10] But, following Chomsky's injunction that we contribute to solving the learnability problem, we would do well to accept as the null hypothesis that they are in fact organized along the same guidelines. This is because doing so is a step towards shrinking the space of grammars through which the learning device has to search. Here, then, is an illustration of how explanatory adequacy can help guide the inquiry. It provides a way to choose among a range of descriptively adequate grammars a 'null hypothesis': the one to start with.

If we give all phrases the shape that NPs, VPs, etc. do, we end up with a family of substitution classes like that below.

(113)
$\text{CP} \rightarrow \overline{\text{C}}$ $\qquad$ $\text{IP} \rightarrow \left\{ \begin{array}{c} \text{(NP)} \\ \text{(CP)} \end{array} \right\} \overline{\text{I}}$

$\overline{\text{C}} \rightarrow \text{C IP}$ $\qquad$ $\overline{\text{I}} \rightarrow \text{I VP}$

$\text{NP} \rightarrow \left\{ \begin{array}{c} \text{(DetP)} \\ \text{(NP's)} \end{array} \right\} \overline{\text{N}}$ $\qquad$ $\text{VP} \rightarrow \overline{\text{V}}$

$\overline{\text{N}} \rightarrow \text{AP } \overline{\text{N}}$ $\qquad$ $\overline{\text{V}} \rightarrow \text{AdvP } \overline{\text{V}}$

10 We will eventually see that PP and DP do, but it requires more exotic constructions than we are now prepared for.

$$\overline{\text{N}} \rightarrow \overline{\text{N}} \left\{ \begin{array}{c} \text{PP} \\ \text{AP} \\ \text{CP} \end{array} \right\} \qquad \overline{\text{V}} \rightarrow \overline{\text{V}} \left\{ \begin{array}{c} \text{PP} \\ \text{AdvP} \\ \text{CP} \end{array} \right\}$$

$$\overline{\text{N}} \rightarrow \text{N (PP) (CP)} \qquad \overline{\text{V}} \rightarrow \text{V (NP) (PP)} \left\{ \begin{array}{c} \text{(IP)} \\ \text{(AP)} \\ \text{(CP)} \end{array} \right\}$$

$$\text{AP} \rightarrow \overline{\text{A}} \qquad \text{AdvP} \rightarrow \overline{\text{Adv}}$$

$$\overline{\text{A}} \rightarrow \overline{\text{A}}\ \text{PP} \qquad \overline{\text{Adv}} \rightarrow \text{DegP}\ \overline{\text{Adv}}$$

$$\overline{\text{A}} \rightarrow \text{DegP}\ \overline{\text{A}} \qquad \overline{\text{Adv}} \rightarrow \text{Adv}$$

$$\overline{\text{A}} \rightarrow \text{A (PP)} \left\{ \begin{array}{c} \text{(CP)} \\ \text{(IP)} \end{array} \right\}$$

$$\text{PP} \rightarrow \overline{\text{P}} \qquad \text{DetP} \rightarrow \overline{\text{Det}}$$

$$\overline{\text{P}} \rightarrow \text{P} \left\{ \begin{array}{c} \text{(NP)} \\ \text{(PP)} \end{array} \right\} \qquad \overline{\text{Det}} \rightarrow \text{Det}$$

These all conform to the following shapes.[11]

(114) $\overline{\text{X}}$ Skeleton: $\text{XP} \rightarrow \text{(ZP)}\ \overline{\text{X}}$
$\overline{\text{X}} \rightarrow \text{QP}\ \overline{\text{X}}$
$\overline{\text{X}} \rightarrow \overline{\text{X}}\ \text{WP}$
$\overline{\text{X}} \rightarrow \text{X (YP) (UP)}$

ZP is called the *Specifier* of XP, WP, QP are called *Adjunct*(s), and YP and UP are called the *Complements* of X.[12]

It should be said that these rules leave out considerable detail. In particular, there are a wide range of things that can stand in adjunct position which are not indicated in these rules. For example, $\overline{\text{V}}$ can have an AP adjoined to it, as in (115).

(115) Sandy [$_{\text{VP}}$ [$_{\text{VP}}$ [$_{\text{VP}}$ saw a man] today] [$_{\text{AP}}$ angry at her]].

11 Except the $\overline{\text{V}}$ rule that introduces verbs. This rule allows for there to be three phrases within the $\overline{\text{V}}$. In fact, however, it appears that at most two phrases are possible. The full three V+XP+YP+ZP sequence allowed by this rule never actually arises.

12 This terminology comes from Jackendoff (1977), as do most of the basic ideas sketched in this section about how phrases are organized. Jackendoff's work is a detailed working out, and considerable extension, of proposals in Chomsky (1970), where there is also a proposal about the underlying building blocks of morpho-syntactic category. Chomsky's work has its seeds in Harris (1946).

And, as noted earlier, an $\overline{\text{N}}$ can have certain kinds of CPs adjoined to them; see (116).

(116) the [$_{\overline{\text{N}}}$ [$_{\overline{\text{N}}}$ [$_{\overline{\text{N}}}$ book] [$_{\text{CP}}$ that Mary read]] [$_{\text{CP}}$ which no one will admit writing]]

If you have a native English speaker handy, it won't take much time to discover many other kinds of combinations that have been left out. I will continue to leave out this detail, invoking it where necessary as we go along.

2.4 Arguments and Modifiers

Charting substitution classes in the way we've been doing characterizes grammatical arrangements of words solely in terms of the words' categorial status. It throws strings of words together according to their status as nouns, verbs, adjectives and the like and ignores all other properties of the particular verbs, nouns, adjectives, etc. chosen. It forms some pretty odd sentences:

(117) a. Jerry danced with pickles.
b. Jerry danced at noon at midnight.
c. Jerry slowly stood still.
d. a green idea

These are odd because the meanings they deliver are so bizarre. But they are still recognizable as grammatical strings of words.

But some combinations which these rules allow seem to go bad in a very different way; consider (118).

(118) a. Jerry laughed Mary.
b. Sam gave it at Jill.
c. Sally died that you should eat better.
d. Jim claimed to Kris.
e. Jerry slapped.

These don't go together into weird meanings; they don't go together at all.

What's wrong, here, is that we've matched up verbs with the material that follows them incorrectly. As can be seen by comparing (118) with (119), if the verbs are followed by different material, the results are grammatical.

(119) a. Jerry laughed.

b. Sam gave it to Kris.
c. Sally died.
d. Jim claimed that you should eat better.
e. Jerry slapped his thigh.

Here, then, is something more particularly about the words themselves that seems to be relevant to the procedure that recognizes grammatical strings. To capture what goes on here, we must do more than know what the category of the words being combined is.

There's another respect in which the particular choices of words seems to play a role in the syntax. Consider the different semantic contributions the NP *Tuesday* makes in (120).

(120) a. I danced Tuesday.
b. I remember Tuesday.

In the first case, *Tuesday* says when *I danced* happened. We say in this case that *Tuesday* is a modifier. It modifies the sentence's meaning by restricting the events denoted by *I danced* to just those that transpire on Tuesday. But this is not the role it has in the second case. Here *Tuesday* refers to the thing remembered. We say in this case that it is an argument of the relation that *remember* denotes.

A similar contrast can be seen in the pair in (121).

(121) a. I kissed her on the bus.
b. I put her on the bus.

Again, *on the bus* is a modifier in the first case. It locates the event described by *I kissed her*; it indicates that this event took place on board the bus. In the second case, by contrast, it names a locations related to the referents of *I* and *her* by *put*. It is an argument.

The semantic role an argument has in a sentence is determined by the word for which it is an argument. The meaning that modifiers contribute to the sentence they're part of is considerably more constant.

There's a way of talking about argumenthood that is commonplace, and which we inherit from Gruber (1965). Gruber, and Fillmore (1968) in a similar paper, was concerned with the problem of verb meanings, and in particular with finding a theory that restricted the kinds of argument types that verbs permit. He speculated that there was a finite, in fact quite small, set of argument types, or 'roles', that could be put together by verbal meanings. He argued

that the roles which verbs combined were always ones that have to do with the metaphysics of motion. For example, a verb like *send* involves three terms, one that can be seen as indicating the Source of the motion, another that denotes the moved term and a third that names the Goal, or endpoint, of that motion. Gruber called the role borne by the term undergoing motion 'Theme.'

(122) Sandy sent his book to Sean.
Source *Theme* *Goal*

On Gruber's thesis, then, verbs are just words that name different kinds of motion. The arguments of verbs will therefore always be the objects, locations, and manners of movement that are involved in motion. Arguments will have roles in the motion relation named by a verb, and what those roles are will be determined by the verb. Those roles Gruber called "theta roles," after the Theme role. You will often see "theta role" expressed as "θ-role." We say, then, that verbs assign θ-roles to their arguments.

There are cases where it is not obvious that motion underlies the relation named by a verb. Nothing has to move in the events described by (123), for example.

(123) a. Sandy showed his book to Sean.
b. Sandy pleases Sean.

But in these cases, one might imagine that there is a kind of abstract motion involved. The image of *his book* might be transferred from Sandy to Sean in (123a), for example; and pleasure might be thought to have moved from Sandy to Sean in (123b).

But there are cases which even metaphorical extensions of the logic of motion look unlikely to characterize. (124) are some.

(124) Sandy finds Sean unpleasant.
Sandy is unhappy.
'Londres' refers to London in Portuguese.
This means that something is wrong.
Sean became unhappy.
They know something.

This way of constraining the meanings of verbs has been pretty firmly abandoned, I believe, as a consequence. In its place, a method has been pursued that tries to see the meanings verbs have as the result of combining a small number

of elemental predicates, like CAUSE, MOVE, and BECOME. On this conception, the roles that arguments play in a sentence can be viewed as resulting from being the arguments of these more basic predicates. We will continue to talk about verbs and the θ-roles they assign.

The relation between verbs and their arguments that is expressed by θ-roles can be seen as a special instance of a more general relationship which goes under the name "selection," or sometimes "s-selection" (for "semantic" selection). This refers to the connection between a verb's (or other similar term's) meaning and the semantic value that its arguments deliver. θ-roles express a similar function: they name the meaning that an argument's semantic value must be compatible with. But the relation holds for other cases too, where the language of θ-roles does not so easily extend. One of those places is where verbs connect with clauses of various types. So, a verb's meaning determines somehow whether the clause it combines with must have the meaning of an interrogative or a declarative, for example.

(125) a. Martha denied that John has left.
b. Martha said that John has left.
c. * Martha wonders that John has left.

(126) a. * Martha denied whether John has left.
b. Martha said whether John has left.
c. Martha wonders whether John has left.

We say of these cases that verbs select or s-select a question or declarative. Note that some verbs are compatible with either, as is *say*.

Though it is hard to see these differences as fitting the functions that θ-roles typically name, I will use the language of θ-roles to describe this relation too.

Now that we have in view this distinction between arguments and non-arguments, let's return to the contrast between (118) and (119), repeated below.

(118) a. Jerry laughed Mary.
b. Sam gave it at Jill.
c. Sally died that you should eat better.
d. Jim claimed to Kris.
e. Jerry slapped.

(119) a. Jerry laughed.

b. Sam gave it to Kris.
c. Sally died.
d. Jim claimed that you should eat better.
e. Jerry slapped his thigh.

In all these examples, I've chosen phrases that cannot get an interpretation as a modifier. The contrast between (118) and (119), then, hinges on giving the verbs involved the right argument. The post-verbal phrases differ with respect to their category. What we discover from these examples is that verbs are picky with respect to what category their arguments belong to. This can also be seen in examples, such as those in (127) and (128), where the difference in the category of an argument is not tangled up so much with the meaning of that argument.

(127) a. Jerry pleases Mary.
b. * Jerry pleases to Mary.

(128) a. * Jerry talks Mary.
b. Jerry talks to Mary.

We say that verbs are *subcategorized* by the category of their argument. Or — this term has been relexicalized — that verbs subcategorize their arguments. Sometimes this is also described as a verb *c-selecting* its argument.[13]

Jackendoff (1977) argues that arguments (when they follow the head they are an argument of) are necessarily in complement position. This is supported by contrasts like:

(129) a. Although Sally didn't Δ Tuesday, she will dance Monday.
b. * Although Sally didn't Δ Tuesday, she will remember Monday.

(130) a. Although Sally won't Δ on the bus, she will kiss her in the car.
b. * Although Sally won't Δ on the bus, she will put her in the car.

Because the phrase following the verb is an argument in (129b) and (130b), it must be within the $\overline{\text{V}}$ which elides, whereas in (129a) and (130c), the phrase following the verb is a modifier and can therefore remain outside the ellipsis.

Jackendoff's thesis is also supported by similar contrasts involving *do so* anaphora, which, like $\overline{\text{V}}$ Ellipsis, finds $\overline{\text{V}}$s.

(131) a. Sam talked to Mary on Tuesday, and Sally *did so* on Thursday.

13 The terms *s-select* and *c-select* come from Pesetsky (1982).

b. Gerry eats chocolate after dinner, and Sandy *does so* before lunch.

(132) a. ?* Sam talked to Mary and Sally *did so* to George

b. * Gerry eats chocolate, and Sandy *does so* marzipan.

c. * Mag proved that she loved chocolate, and Holly *did so* that she loved marzipan.

If we examine the positions that *do so* may stand in, we will find that it has the same distribution as $\overline{V}$s: it may appear between a subject NP and a sentence final PP, as in (131). In this respect, then, *do so* is like *one* in that it is a lexical expression of a certain phrase. If this is granted, then the ungrammaticality of the examples in (132) indicates that the material following the verb in these cases must be within the smallest $\overline{V}$. This would explain why this material cannot be positioned outside of *do so*. And what distinguishes the cases in (131) and (132) is that the phrases following *do so* in (132) are arguments, whereas those in (131) aren't.

To the extent that we can tell, the same is true with respect to arguments of other classes of words. It's difficult to be certain of the argument-status of terms which combine with nouns, for instance, but Jackendoff's claim seems correct here as well.[14]

(133) a. I'll listen to your long, careful discussion of it, if you'll listen to my short *one*.
one = "careful discussion of it"

b. * I'll listen to your long, careful discussion of it, if you'll listen to my short *one* of it.
one = "careful discussion"

c. I'll listen to your long, careful discussion in class, if you'll listen to my short *one* in the office.
one = "careful discussion"

The contrast between (133a) and (133b) will follow if *of it* must be positioned within the smallest $\overline{N}$. The contrast between (133b) and (133c) corresponds to the differing argument-status of the PPs involved: *of it* is more strongly perceived as an argument of *discussion* than is *in class*. As with the *do so* and $\overline{V}$

14 There is evidence that nouns assign θ-roles only when they have a verb-like use; that is, when they are used to describe processes or events, and not when they are used to name things (see Grimshaw 1990). The examples in (133) are constructed with this in mind. We return to this aspect of nouns in a later chapter.

Ellipsis facts, then, this contrast supports the hypothesis that arguments and modifiers are fit into phrases in different positions.[15]

Okay, to summarize: we're looking for a way to factor into our procedure for recognizing grammatical sentences enough of the meanings of the words involved to guarantee that Verbs and Nouns (and perhaps other words) combine with the arguments they select and subcategorize. Moreover, when these arguments follow them, we must find a way of guaranteeing that they are in the non-recursive $\overline{\text{X}}$: the complement position.

We can ensure that these arguments are in the non-recursive part of the $\overline{\text{X}}$ if we force them to bear a θ-role, and allow θ-roles to be assigned only to complement positions. We need also to describe the fact that when a verb has a θ-role, there must be an argument present in the syntax which bears that θ-role. It is customary to divide this task into two parts, which can be expressed as follows:[16]

(134) THE THETA CRITERION

a. For every θ-role there is a position to which that θ-role is assigned.

b. For every θ-position, there is something with an appropriate semantic value that occupies that position (i.e., the argument).

It is usual to strengthen the Theta Criterion to a bijection, because of cases like (135).

(135) Sally showed John *doesn't mean* Sally showed John himself.

Without constraining the THETA CRITERION to a bijection, we might expect (135) to get such an interpretation since presumably the NP *John* could name the object which bears both the Theme and Goal θ-roles. So we change this to (136).[17]

(136) THE THETA CRITERION

15 Baker (1978) is, perhaps, the first to argue from *one* anaphora for this conclusion about where arguments are positioned within NPs.

16 The Theta Criterion comes from Chomsky's important *Lectures on Government and Binding*.

17 The THETA CRITERION is also often formulated in terms of a bijective relation between θ-roles, or θ-positions, and arguments. That is, it is sometimes written to say: "For every θ-role (or positions) there is exactly one argument and for every argument there is exactly one θ-role (or position). (In *Lectures on Government and Binding* it is formulated in various ways, including these two.) The difference between this alternative formulation and the one I've given here is that mine does not force every argument to receive a θ-role, whereas the alternative does. I've decided to place this requirement in another principle, which we'll come to shortly.

a. For every θ-role there is exactly one position to which that θ-role is assigned.
b. For every θ-position, there is exactly one thing with an appropriate semantic value that occupies that position (i.e., the argument).

When we add to this the fact that verbs also specify the categories of their arguments, we get something like (137).

(137) For every θ-role assigned by some X, X c-selects the phrase that bears that θ-role.

We need to worry about cases like the following, of course, in which there appears to be an optional argument.

(138) a. Martha ate (pie).
b. It seems (to me) that Marty left.

Here we might imagine either that there actually is an object in these cases that bears the θ-role, but that argument is unpronounced. Alternatively, we might conjecture that something relaxes the condition which forces every θ-role to be assigned to a position holding an argument. The common wisdom is that both possibilities exist — we will return to this issue in some detail later. For now, let us imagine that there is a lexically determined process which allows θ-roles for certain predicates to not be assigned.

Some have suggested that (134b) and (137) should be collapsed, and in particular, that there is a means by which the categorial type of some argument can be determined from its θ-role.[18] Grimshaw (1979) provides a way of viewing this hypothesis which has gained some popularity. Her idea is that one of the functions that makes up the learning device assigns a categorial status to arguments on the basis of their θ-role. She calls this function "Canonical Structural Realization" (CSR). She sketches how this function might work by way of examples that compare CPs with NPs.

So let's look as some of the facts she considers. Note first that CPs may distinguish themselves as according to whether they denote Propositions, Exclamatives or Questions. Let's use the language of θ-roles to describe these types of clauses. These θ-roles can sometimes be borne by NPs too:

(139) a. John asked me $\left\{ \begin{array}{l} \text{what the time is} \\ \text{the time} \end{array} \right.$ (Question)

18 The reason c-selection is usually thought to be derivable from s-selection, rather than the other way round is tied to Chomsky's "epistemological priority" argument, see Pesetsky (1982).

b. I'll assume { that he's intelligent / his intelligence } (Proposition)

c. Bill couldn't believe { how hot it is / the heat } (Question)

In these cases, then, the verbs s-select either Q(uestion), P(roposition) or E(xclamative) and c-select either an NP or CP.

There are other verbs, however, which s-select these very same θ-roles, but c-select only CPs.

(140) a. John wondered { what the time was / *the time } (Question)

b. I'll pretend { that he's intelligent / *his intelligence } (Proposition)

c. Bill complained { how hot it was / *the heat } (Exclamative)

Here then, we have a special instance of the difference in s-selection and c-selection that needs to be overcome if one is to be derived from the other.

Grimshaw's suggestion is that the CSR of Questions, Propositions and Exclamatives is CP and that those verbs which allow these θ-roles to be borne by NPs are learned on a case-by-case basis. Thus, this is a partial collapse of c-selection to s-selection. And it predicts that every verb that s-selects a Q, P or E will c-select a CP; that is, there should be no verbs that express these θ-roles with an NP only. This seems to be correct.

Whether or not this project can be maintained for the situation involving the relation between CPs and NP and the θ-roles they bear, I don't think a parallel story holds for the complements of other categorial type. Moreover, the scheme Grimshaw proposes won't help determine which verbs select non-finite as opposed to finite clauses, which also seems to be a rather language particular fact. The problem we face here is the same problem we face when we try to understand if morphosyntactic word-class can be expressed in solely semantic terms. It remains unknown what the relationship between word-class and semantic type is, and consequently, we are not in a position to yet understand what the connection, if any, is between c-selection and s-selection. From now on let us assume that c-selection is at least in part independent of s-selection, and determined on a verb-by-verb basis.

Interestingly, however, it looks like the thesis that c-selection can be derived from s-selection fares better when external arguments are concerned. To begin

with, the range of categories that serve as external arguments looks somewhat less varied; to a large extent, only NPs and CPs seem to be clausal subjects in English.[19] And second, when a θ-role is consistent with either NP and CP, any kind of CP is possible as is an NP:

(141) { That John left / To have to leave / Leaving / The fact } { bothers Mary. / makes Mary happy. }

By contrast, when the subject θ-role is incompatible with the meanings that CPs yield they are banned from Specifier position:

(142) { ∗That John left / ∗To have to leave / ∗Leaving / John } { kisses Mary. / likes Mary. }

Subject arguments do not seem to be c-selected by the verbs involved. The θ-role they bear is sufficient to determine what category they can be. Only complements are c-selected. This will require weakening (137) to something like (143).

(143) a. A word c-selects its complements.
b. If a word c-selects Y, then it θ-marks Y.

We can summarize what we've discovered so far with (144).

(144) a. If a verb has a θ-role, then there is exactly one syntactic position to which that θ-role is assigned.
b. A θ-marked position must be occupied by something with the appropriate semantic value.
c. A verb c-selects its complements.
d. If α c-selects β then α s-selection β (aka θ-role assignment).

The statements in (144a) and (144b) are the Theta Criterion, whereas those in (144c) and (144d) concern the relation between c-selection and s-selection which we've just reviewed. The Theta Criterion insists that for every θ-role that some term has, there will be a unique position occupied by an argument

19 With the exception of cases like "Under the bed is a slipper," plausibly instances of impersonal constructions with inversion; see Stowell (1981) and Rochemont and Culicover (1990).

in the sentence holding that verb. (144c) and (144d) determine whether that argument will be c-selected or not.

To force arguments to be within the smallest $\overline{\text{X}}$, it will now be sufficient to force the θ-position for that argument to be within the smallest $\overline{\text{X}}$. We want this effect for complement arguments only — we don't want to force "subject" arguments into $\overline{\text{X}}$ — so one way of doing this would be to restrict those positions that are c-selected to just those within the smallest $\overline{\text{X}}$. This would mean that we'd have two principles: one that determines the c-selected position for verbs, and another, yet to be determined, which locates the s-selected position for subjects. We're going to see, however, that the procedure for locating the θ-positions for both subject and object arguments is the same, or very nearly so, and so we won't take precisely this course.

Instead, we will follow a popular view of these principles that is first found in Chomsky's *Lectures on Government and Binding*. He formulates there what he calls "The Projection Principle," which is responsible for mapping the argument structure of a verb — or head more generally — into a syntactic representation. I will formulate his principle as (145).[20]

(145) The Projection Principle

i. For α, a position, if α is a sister to X^0,[21] then X^0 c-selects α's contents.

ii. If α s-selects β, then α and β are sisters.

The second part of the Projection Principle does what we are in search of. It forces arguments of a verb to be in the lowest $\overline{\text{X}}$, for only in that position will it be a sister to the verb.[22] Note that this principle is not restricted to verbs and their projections, it spreads what we've discovered about VPs to all other categories. This, so far as I know, is correct.

As presently formulated, the second part of the Projection Principle wrongly forces subjects into the smallest $\overline{\text{V}}$ of the verb that assigns it a θ-role. We will see, however, that this problem is only apparent. Once we discover what is truly responsible for assigning the subject its θ-role, this problem is resolved (or, depending on how certain particulars play out, mutated into a different problem).

20 Chomsky's own formulation builds in various other properties that we will encounter later on; see in particular the discussion in Chomsky (1981, pp. 34–48).

21 "X^0" is the position in a phrase marker into which a word is inserted. It is a way of explicitly referencing the (smallest) head of a phrase.

22 X and Y are *sisters* if every phrase including one includes the other.

The first part of the Projection Principle is just (144c). It has the interesting consequence of preventing non-arguments from standing in the smallest $\overline{X}$. Thus, the Projection Principle not only has the effect of forcing arguments into the smallest $\overline{X}$, but also of forcing non-arguments out of this position. Whether this stronger result is correct is rather difficult to determine. We will eventually examine phenomena that might bear on it.

2.5 Deriving Phrase Structure Rules

An interesting consequence of the Projection Principle is that it factors into the lexical specification of the verbs everything needed to know what sort of phrases will be found in the non-recursive part of $\overline{X}$s. Tim Stowell, in his 1981 MIT dissertation, argues that the information about what resides in the lowest $\overline{V}$ should only be found as part of the verb's lexical content.[23] This sort of information doesn't properly reside in the phrase structure rules, since it is information that is tied to the particular choice of word and not the pure form that sentences may take. In fact, the information phrase structure rules give about the contents of the smallest $\overline{X}$ can now be seen as merely a summation of what is possible across particular lexical items filling the head slot. Thus, we should factor out of the phrase structure rules information which concerns the categorial nature of the complements involved. We can do the same for the subject arguments as well, since their categorial nature is derived from their θ-role.

What we have seen, then, is that the phrase structure rules can be stripped of a great deal of their information. Indeed, what is left is largely what the $\overline{X}$ skeleton expresses and the categorial specification of non-arguments. There is some hope, I think, for the view that the categorial specifications for non-arguments will follow entirely from the meanings that categorial types may have. So it might be that the fact that $\overline{V}$ can combine with non-argument PPs, AdvPs, CPs and NPs may follow entirely from the meanings that categories of these types may carry. Thus, CPs may denote "reasons" (say, as in *because* clauses) and AdvPs can denote manners, and PPs can denote locations or times, as can NPs, and these are just the sorts of things that allow for combination with $\overline{V}$s to form other $\overline{V}$s. Similarly, it might be that $\overline{N}$s may only combine with the types that PPs, AdjPs and CPs belong to because these are the only types that, once joined with $\overline{N}$s, produce another $\overline{N}$. Let us suppose that this is the case.

23 See Stowell (1981).

(Note that the range of categories possible in these positions is relatively free, suggesting that there are few, if any, constraints on category type.) Since the inventory of categories varies from language to language we might, just to be safe, factor this information out of the phrase structure rules into a language particular set of statements of the form in (146).

(146) a. If α modifies $\overline{\text{N}}$, then α must be ...
b. If α modifies $\overline{\text{V}}$, then α must be ...
⋮

The "..." will carry lists of category types.

If this project is successful, then the Phrase Structure rules of English collapse in full to the $\overline{\text{X}}$ Skeleton. Some have argued for a picture of language variation that makes the hierarchical arrangements of constituents that the $\overline{\text{X}}$ Skeleton, together with the PROJECTION PRINCIPLE and THETA CRITERION and whatever yields (146), completely immutable.[24] All that varies across languages is the linear order in which the terms that follow the arrows in the $\overline{\text{X}}$ Skeleton may have. So, the phrase structure component of the grammar might have nothing more than (147) in it, where "$\{\alpha,\beta\}$" should be understood as representing both the string α+β and the string β+α.

(147) a. $\text{XP} \rightarrow \{(\alpha\text{P}), \overline{\text{X}}\}$
b. $\overline{\text{X}} \rightarrow \{\overline{\text{X}}, \beta\text{P}\}$
c. $\overline{\text{X}} \rightarrow \{\text{X}^0, (\gamma\text{P}), (\psi\text{P})\}$

What the morpho-syntactic category of α, β and γ are is fully determined by the c-selection properties of X^0 and the language particular principles governing modifier types (i.e., (146)).

Note, incidentally, that I've somewhat arbitrarily set the highest number of things that can share the head with the smallest $\overline{\text{X}}$ to two. That is, I've decided that there can be at most two complements. That seems to be supported for verbs and adjectives, when can have at the same time a PP and CP complement, for instance, but no combination of complements that goes beyond two. In later chapters we will revisit this issue.

The linear arrangements of these constituents must then be determined by the language particular part of the grammar. There is evidence from language typology that whatever it is that determines the order of Specifier and $\overline{\text{X}}$ is

24 See Travis (1984), for example.

independent of what determines the order of heads and their complements. There is no widely agreed upon account of what is responsible for this factor, so let's leave this for the future. There is also typological evidence[25] that the order or complement and head correlates with the order of modifier and $\overline{X}$. One way of expressing this is with a "headedness parameter" which specifies whether the head of a phrase may come initially or finally in its (immediate) projection. This predicts that complements will either all precede or follow their heads, and not come among them. While this is not superficially true (German/Dutch, for example, don't seem to obey this), it does look like this could be true of the underlying arrangements of these constituents and it does seem to capture a tendency that languages have.

As a starting point, then, let's take the view that languages linearize their phrasal constituents by way of setting separately the linear order of the immediate constituents of XP and $\overline{X}$. This can be achieved by letting languages pick the values "first" and "last" for the terms in (148).

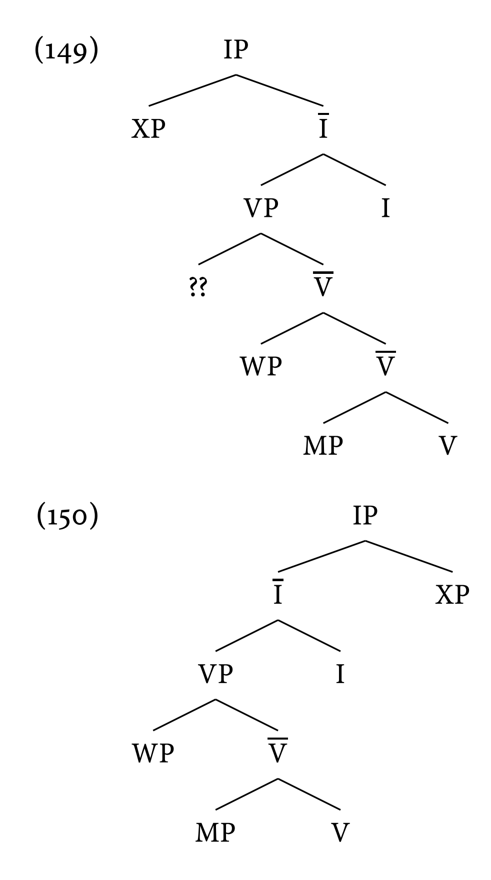
(148) If σ is an immediate daughter of π, π an $\overline{X}$ projection[26] of α, then choose a value in "< >" for:

a. σ = Specifier of π: <first in π, last in π>

b. σ = the term π is an $\overline{X}$ Projection of: <first in π, last in π>, modulo (148a)

This connects the linear order of head to complement with the linear order of head to adjunct, which Dryer's work suggests might be correct.[27] So, for instance, a language that sets Specifier to "first" and Projection-of-α to "last" will license sentence structures like those in (149), whereas a language that sets both terms to "last" will produce structures like those in (150).

25 See Greenberg (1963) and Dryer (1992).
26 β is an "$\overline{X}$ Projection of α" iff β dominates α and they have the same head.
27 See also Saito (1985) and Saito and Fukui (1998).

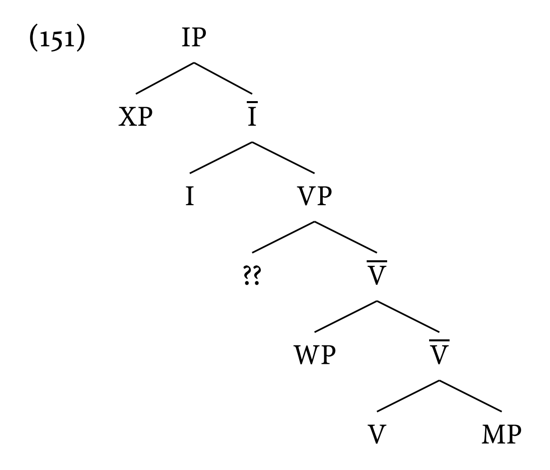


The categorial values for MP will be determined by the c-selection specification of the verb involved. The categorial values for XP will be determined by the θ-role it receives. And the categorial values for WP will be whatever (147) for the language in question allows to modify $\overline{\text{V}}$s. We haven't yet discovered what sits in the Specifier of VP, so this spot is marked with "??."

The linearization parameters in (148) produce these phrase markers in the following way. Setting Specifier to "first" in (149) linearizes XP and ?? so that they precede $\overline{\text{I}}$ and $\overline{\text{V}}$ respectively. Setting Projection-of-α to "last" makes every other $\overline{\text{I}}$ and $\overline{\text{V}}$, as well as I and V, follow the phrase they are sisters to. As a consequence WP, MP and VP precede the phrase they are complements to or modifiers of. In (150), by contrast, Specifier is set to "last," which linearizes XP and ?? so that they follow $\overline{\text{I}}$ and $\overline{\text{V}}$ respectively. As with (149), Projection-of-α is set to "last" in (150) and the consequence for the position of WP, MP and VP is the same.

Restricting the linearization options to just those in (148) blocks certain phrase markers. It blocks languages, for instance, in which the complement

to a verb falls on a different side of that verb than does a complement to a noun (or any other category). That is, it forces languages to unify the linearization of Specifier, Complement and modifier across phrase types. It is not hard to find languages that seem to violate this restriction, but as Greenberg and Dryer find, there is a tendency for languages to avoid this type. Similarly, (148) prevents all languages that put modifiers to one side of the $\bar{X}$ they modify but put complements to the other side. For instance, phrase markers like (151) are prevented.

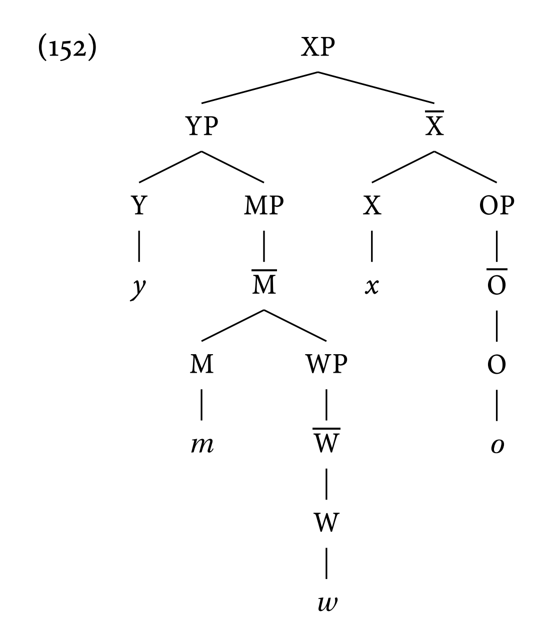


This phrase marker linearizes $\bar{V}$ (a projection of V) "last" relative to WP, but linearizes V (also a projection of V) "first" relative to its complement. Clearly there are languages of this unexpected type; English seems to look precisely like (151).

This proposal, then, seems clearly too restrictive. Nonetheless, it will be our starting point. In the chapters that follow we will explore ways of loosening this model so that it is enabled to account for the range of language types we do see without losing the trends in linear organization that Greenberg and Dryer have discovered. What we have now is not yet complete enough to really engage this problem.

In fact, the linearization scheme in (148) is itself not yet complete enough to generate the strings we want to associate with the phrase markers it allows, for example those in (149) and (150). All (148) does is linearize the phrases within a sentence. It does not determine how the strings of words within those phrases are linearized relative to the other phrases. To see this, consider a phrase marker like that in (152) on the following page, in which lower-cased letters should be understood as representing words. This phrase marker arises

(152)



by imposing the $\overline{\text{X}}$ Skeleton and setting Specifier to "first" and Projection-of-α also to "first." What we would like is for this to be sufficient to generate the string *ymwxo*. Instead, however, all that these settings give us is the information in (153).

(153) a. Y precedes MP
b. M precedes WP
c. YP precedes $\overline{\text{X}}$
d. X precedes OP

What's required is something to determine how the information in (153) determines the linear order of y relative to the words within MP and $\overline{\text{X}}$, the linear order of m relative to the words in WP, and the linear order of x relative to the words within OP. Let's turn to that now.[28]

Recall that in defining morpho-syntactic category, we entertained the hypothesis that looking at only adjacent terms would be sufficient for defining the relevant substitution classes. As it happens, in defining phrases we have also obeyed this constraint. As a result, phrases are always strings of adjacent terms. Let's elevate this too to an hypothesis:

(154) CONTIGUITY
Let $\tau = \{ \alpha_1, \alpha_2, \ldots, \alpha_n \}$ be terminals dominated by ζ. The string formed from τ cannot contain β if β is not dominated by ζ.

If Contiguity holds, it is possible to determine from (153) what the consequent linearization for all the words in (152) is. If the words in MP must be adjacent to each other, then (153a) is enough to know that y precedes all those words (i.e. m and w). Similarly, if the all the words in YP must form a contiguous string, and all the words in $\overline{\text{X}}$ must too, then from (153c) it is possible to deduce that every word in YP (= y, m and w) must precede every word in $\overline{\text{X}}$ (= x and o). All that is required is an explicit statement that the words within a phrase, α, are linearized with respect to the words in phrase β in the same way that α is linearized to β. This can be done with (155).

(155) $\alpha < \beta =_{\text{def.}} \alpha$ precedes β.
$\{_\delta\ \alpha, \beta\} =_{\text{def.}}$ $\begin{array}{c}\delta\\ \diagup\diagdown\\ \alpha\quad\beta\end{array}$ or $\begin{array}{c}\delta\\ \diagup\diagdown\\ \beta\quad\alpha\end{array}$.[29]

a. For all words, x and y, within a phrase marker, either $x < y$ or $y < x$.

28 The relationship between phrase markers and the strings they map onto is investigated in a vast literature, where a wide array of different formalizations can be found. Two widely read classics are Chomsky (1985) and Lasnik and Kupin (1977). For a simple introduction to the formalisms of graph theory (from which phrase markers are drawn) and their projection onto strings, see Partee, Meulen, and Wall (1990).

29 α and β are said to be "sisters" in this case, and δ is called α and β's "mother."

b. Let X and Y be points on a phrase marker. If X < Y, then $x < y$ for all x dominated by X, and all y dominated by Y.

(155a) merely makes explicit that all the words in a phrase marker must have a linear relation to every other word in a phrase marker. (155b) determines how these linear relations are derived from the language particular orderings imposed upon phrases. It also derives Contiguity.

On this view, then, the phrase structures of languages are the result of four fixed universals — the $\overline{\text{X}}$ Skeleton, the Theta Criterion, the Projection Principle, and the linearization principles in (155) — plus the language particular pieces of information in (156).

(156) a. Specifying the categories of modifiers (i.e., (146)).
b. Setting the "headedness parameter" (i.e., (148)).
c. A vocabulary of lexical items s-select and c-select arguments.

There's a sense, then, in which languages do not actually have Phrase Structure rules. They are merely the epiphenomena that emerge when the various factors of Universal Grammar and language particular information are combined. This theory, if correct, meets the criterion of explanatory adequacy. It provides both inviolable constraints (i.e., $\overline{\text{X}}$ Theory, the Theta Criterion, the Projection Principle and (155)) and an evaluation metric (i.e., (156) and the language particular vagaries of vocabulary).

Notice how the evaluation metric this proposal embraces is quite different from the "simplicity" metric suggested in Chomsky's early work. The evaluation metric here involves learning the word-by-word selection requirements and fixing parameter values in the headedness linearization procedure in (156). This proposal has the following form: inviolable constraints come in the form of immutable principles, while the evaluation metric (once lexical idiosyncrasies are removed) consists of principles with a menu of parameters that are set on a language particular basis. Theories that have this general form are said to belong to the "Principles and Parameters" framework. This conception of what explanatory grammars might look like was suggested by Noam Chomsky and his collaborators in the late 1970's, and much of the work of the 80's and early 90's has this form. In 1981, Chomsky published an ambitious book in which he organized much of the work of that time into a principles and parameters form. This book, *Lectures on Government and Binding*, serves as a rough starting point for much of my exposition in these lectures.

In moving from a battery of English specific phrase structure rules to the more explanatory interaction between $\overline{\text{X}}$ Theory and the c-selection requirements of verbs, language particular settings of modifier types, etc., we have lost some information. Because that transition removed any reference to categories, it is no longer possible to order complements in situations, like (157), where there are more than two.

(157) Sheila put this on the table.
compare: *Sheila put on the table this.

This information was conveyed in the phrase structure rules by way of referencing category type. We had rules such as (158) for instance which ensure that if a verb is followed by two complements, the first will be the NP.

(158) $\overline{\text{V}} \rightarrow$ V (NP) (PP) (CP)

Because the $\overline{\text{X}}$ Skeleton does not have information about category type in it, it is not possible to use the $\overline{\text{X}}$ Skeleton to order complements. Nor would we want to rely on the c-selection requirements of verbs to do this. That would amount to the claim that the order of complements varies as a function of the verb involved. But the fact of English is that no matter what verb is selected, the complements line up in the way that (158) requires.

There must be another component of the grammar which expresses this information. Indeed, we have several outstanding problems concerning the position that arguments take. For instance, we have the paradoxical behavior of CP objects yet to understand. How are object CPs at once within the lowest $\overline{\text{X}}$ with respect to constituency tests like *do so* anaphora, but also linearized so that they appear after modifiers that are not within this $\overline{\text{X}}$? Among the loose-ends fraying at the edges of the proposals presented in this chapter, the correct characterization of argument placement looms large. This will be the subject of the next chapter.

3

Positioning Arguments

Our grammar of English presently consists of a series of principles that determine the shapes that phrase markers may have. Those principles make reference to the morpho-syntactic category that words belong to, and also to the semantic relation that arguments and modifiers have to the things that they combine with. These principles are collected in (1).

(1) a. $\overline{\text{X}}$ Skeleton: XP → {(αP), $\overline{\text{X}}$}

$\overline{\text{X}}$ → {$\overline{\text{X}}$, βP}

$\overline{\text{X}}$ → {X^0, (γP), (ψP)}

b. The Theta Criterion

i. For every θ-role there is exactly one position to which that θ-role is assigned.

ii. For every θ-position, there is exactly one thing with an appropriate semantic value that occupies that position (i.e., the argument).

c. The Projection Principle

i. For α, a position, if α is a sister to X^0, then X^0 c-selects α's contents.

ii. If α s-selects β, then α and β are sisters.

d. If a word c-selects Y, then it θ-marks Y.

e. Modification Parameters
   i. If α modifies $\overline{\text{N}}$, then α must be AP, PP or CP
   ii. If α modifies $\overline{\text{V}}$, then α must be AdvP, PP, CP or NP

⋮

The $\overline{\text{X}}$ Skeleton forces a certain shape on all arrangements of words, grouping them into phrases. The Projection Principle in conjunction with (1d) ensures that θ-positions are sisters to the terms that assign the θ-roles, and that when the term assigning a θ-role is a head, that the argument it assigns a θ-role to is also c-selected. This leaves the problem of correctly determining the subject's θ-position — a problem whose solution we are working towards. The Theta Criterion ensures that for every θ-role associated with some predicate, there will be exactly one θ-position in the syntax, and that this position will be occupied by an argument. This, together with the Projection Principle will correctly place objects deeper than non-arguments within the phrase that contains them. Finally (1e) lists the categorial status that modifiers may have, depending on the term that is being modified.

In addition to these principles, we discussed a method for mapping the groupings of words these principles control into strings of words. In particular, I suggested that there were two linearization parameters that control how the phrases within some other phrase are given a linear order. One of those parameters determines whether Specifiers of phrases come before everything else in the phrase they are Specifiers for, or come after everything else. A second parameter determines whether the phrase that contains the head precedes or follows the phrase it combines with. These parameters are repeated in (2).

(2) If σ is an immediate daughter of π, π an $\overline{\text{X}}$ projection of α, then choose a value in "< >" for:
   a. σ = Specifier of π: <first in π, last in π>
   b. σ = the term π is an $\overline{\text{X}}$ Projection of: <first in π, last in π>, modulo (2a)

These parameters express an hypothesis about how languages can vary. They say, for instance, that the Specifier of IP will always come on the same side of the rest of the material in IP that the Specifiers of every other phrase in that language do. And they predict that languages will put the heads of phrases always to the left or the right of their complements, no matter what category that phrase is. The objects of verbs, in other words, should fall on the same side of the verb that the objects of adjectives, or prepositions, do. I remarked

that these cross-linguistic predictions are somewhat spottily supported. We'll return to this matter in a later chapter and take a closer look at the facts.

The linearization parameters are coupled with a constraint on linearization that enforces Contiguity. That constraint requires that all the words within a phrase, α, follow or precede all the words in a phrase β in accordance with how the linearization parameters arrange α and β. If α precedes β, then every word in α will precede every word in β, and if α follows β, then every word in α will follow every word in β.

These principles are all declarative statements. They express constraints on the syntactic organization that sentences may have. What's missing from this picture is the engine that produces that organization. Nothing in this grammar tells us what gets those words grouped into phrases; we just have a series of hypothesized generalizations about what the resulting groups of words must look like.

I'll briefly sketch now a method of generating these structures. I do this because it'll be handy to have something concrete in front of us when we think about how the representations for sentences are created. I won't give any reason for thinking this particular procedure is correct, and so I don't suggest you believe it. We'll have a chance to examine some of its claims later on. Until then, let us use this as a crutch as we limp from language data to speculations about the constraints those data are evidence for. The procedure I'll sketch is very loosely built upon Chomsky's *Aspects of the Theory of Syntax*.

The first thing we'll need is a device that assembles the phrases that the $\overline{X}$ Skeleton constrains. For that purpose, I will define three functions, all named GROUP, that differ with respect to the number of arguments they take. These functions simply arrange the N^0, $\overline{N}$, VP and other such symbols into phrase markers of the sort that the phrase structure rules from the previous chapter built. These functions are defined in (3) on the next page.

The second thing we'll need is a way of matching the words in a sentence to the positions produced by GROUP in the phrase marker it constructs. This procedure is called LEXICAL INSERTION, and it's spelled out in (4).

Together, GROUP and LEXICAL INSERTION will produce a wide array of phrase markers. The principles in (1) can then be thought of as filters on that set of phrase markers. They will permit only those that arrange the symbols N^0, VP and so on in the ways that the $\overline{X}$ Skeleton permits, and they will arrange arguments and modifiers in ways that conform to the Projection Principle and Theta Criterion. The phrase markers that GROUP and LEXICAL INSERTION form are sometimes called "Deep Structures," or "D-structures." We can

(3) Let δ be any of the symbols X^0, $\overline{X}$, XP, where "X" ranges over the morphosyntactic categories in the language, and let α, β and γ be δ or GROUP(α) or GROUP(α)(β) or GROUP(α)(β)(γ).

a. GROUP(α) $=_{\text{def.}} \{_\delta\ \alpha\}$

b. GROUP(α)(β) $=_{\text{def.}} \{_\delta\ \alpha, \beta\ \}$

c. GROUP(α)(β)(γ) $=_{\text{def.}} \{_\delta\ \alpha, \beta, \gamma\ \}$

(4) Lexical Insertion
Insert into every X^0 in a phrase marker a morpheme of category X.

think of the principles in (1) as well-formedness conditions on D-structures.

There are a variety of problems for the model that we took note of as we developed it. Let's list the more salient of these here so that we do not forget them.

(5) a. For many phrases, the distinction between XP and $\overline{X}$ has not been motivated. We have not seen, for instance, what phrase lies within the Specifiers of VP, AP, PP, AdvP, or CP. For all of these phrases, we have only seen evidence for $\overline{X}$s, a kind of $\overline{X}$ that is recursive and another that introduces complements.

b. The Projection Principle describes the conditions under which θ-roles are assigned, but it only correctly delivers θ-roles to objects. Under what conditions is the θ-role delivered to a subject argument?

c. The CP complements to verbs present a paradox. They appear to be in complement position with respect to tests like *do so* anaphora and $\overline{V}$ Ellipsis, but they are positioned linearly as if they are not in complement position. They seem to appear following modifiers of the verb.

d. Our present system does not place any constraints on the linear order that two complements take with respect to each other. We should expect to find, for instance, that the objects of verbs could come in any linear order. Instead, what we seem to find is that a NP object to a verb precedes a PP object, and that a CP object of a verb follows all other objects of the verb (this is part of the problem in (5c)).

e. The head of an IP can be an inflectional morpheme, and when it is, our rules will position that inflectional morpheme before, and

separated from, all verbs in the sentence. This is wrong. The inflectional morpheme appears instead on an immediately following verb.

We will tackle the first four of these problems in this chapter. The problem in (5e) will have to wait until the following chapter.

Before doing that, however, there is another difficulty that crept into our model when we transitioned from phrase structure rules to the system of constraints on D-structures that we presently have. Our first task will be to look at that difficulty.

3.1 Obligatory Specifiers

The $\overline{\text{X}}$ Skeleton makes all of the phrases within some maximal projection optional. The presence of modifying phrases is, in fact, completely optional. The presence of complements is determined by the existence of θ-roles: if the head of the phrase has θ-roles, then phrases in these positions will be obligatory, forced by the Theta Criterion and the Projection Principle. What about the phrase in Specifier position? In general, the phrases in these positions are optional.

Occasionally, phrases in specifier positions are forced by processes that are not well understood. For instance, the presence of something in Specifier of NP seems to be determined by whether the head noun is singular or plural:

(6) a. I like horses.
    b. I like the horse.
    c. * I like horse.

This might be due to a semantic effect, though there is considerable language variation here whose source is not known. Another possibility is that there is a determiner present even in (6a), but that it is silent. This would allow for the possibility that Specifier of NP is obligatorily filled, accounting, then, for the ungrammaticality of (6c). Let's leave the status of the specifier of NP open, for the moment. We'll have a chance to revisit this issue when we examine more closely the structure of noun phrases.

Phrases in Specifier of IP, however, buck the trend and are always obligatory. This was one of the facts encoded in our original phrase structure rules that has been lost in the transition to the $\overline{\text{X}}$ Skeleton plus constraints system that we are now employing. In cases where the IP contains a subject argument,

the obligatoriness of this subject is plausibly derived in the same way that the obligatoriness of complements is: by the Theta Criterion and the Projection Principle. Somehow or other, the Projection Principle is going to have to be fixed so that it guarantees that there is a θ-position for the subject argument, and the Theta Criterion will force an argument into this position. If the θ-position happens to be specifier of IP, then this will guarantee the presence of something in specifier of IPs whose verbs assign a subject θ-role. But, interestingly, even in IPs whose verbs do not have a subject θ-role, the presence of something in Specifier of IP is obligatory. The verb *seem*, for instance, has only one θ-role, and that is assigned to its clausal complement. And yet, as (7) shows, an IP containing this verb must surface with something in its specifier position.

(7) a. It seems that we are behind.
    b. * Seems that we are behind.

The *it* in (7a) appears to have no meaning whatsoever, and is merely present in order to occupy specifier of IP. It is called an "expletive" or "pleonastic" term, to indicate its lack of semantic content.

To recapture this bit of information, Chomsky proposes in *Lectures on Government and Binding* adding another statement to the Projection Principle which simply requires that specifier of IP be filled. This is known as the extension to the Projection Principle, or much more commonly, the EPP.

(8) Extension of the Projection Principle (EPP)
    The Specifier of IP must have a phrase in it.

In the normal case, when the verb of a sentence has a "subject" θ-role, the Extended Projection Principle will be satisfied by the presence of an argument phrase, whose presence will also satisfy the Theta Criterion. But in the relatively rare case when the verb does not have a subject θ-role, it will still demand the presence of something, and the expletive is invoked as a consequence.

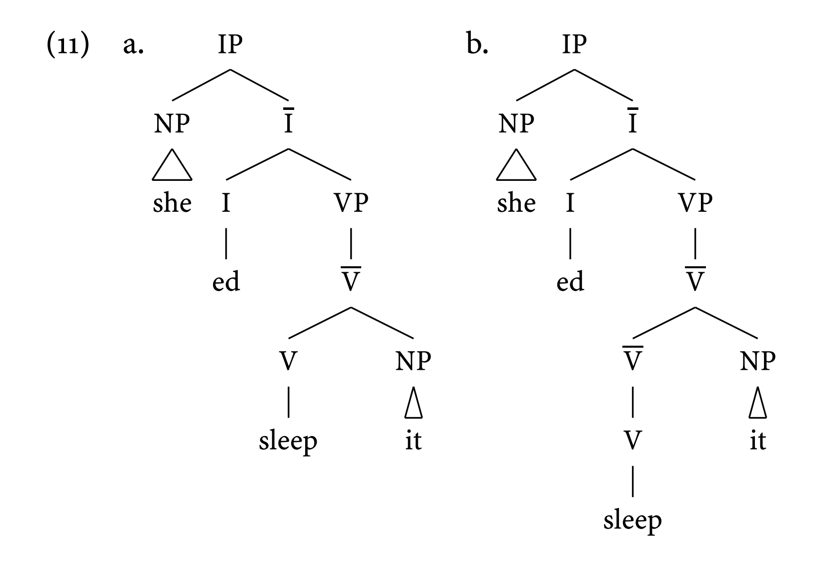
Notice that this system restricts the use of the expletive to just those circumstances where there is no subject θ-role. When the verb of some sentence has a subject θ-role, the Theta Criterion will require that the Specifier of IP be occupied by an argument and expletives, by virtue of their semantic vacuity, cannot function as arguments. This, then, is why (9) does not allow the *it* in specifier of IP to be interpreted as an expletive.

(9) It discovered the problem

In fact, the distribution of expletives is extremely limited. They are found only in Specifier positions to which no θ-role is associated. As we've seen, they're not found in Specifier positions associated with θ-roles, but they are also not found in complement or adjunct positions. Thus an example like (10) is ungrammatical on either of the parses indicated in (11).

(10) *She slept it.

(11)



The parse in (11a) is already blocked by the Projection Principle, the first clause of which requires that things in complement position be c-selected and θ-marked by the neighboring head. This isn't the case in (11a).

I don't know of anything in the literature that is explicitly designed to exclude (11b), so I suggest that something along the lines of (12) is responsible.

(12) $\{_{\overline{X}}\ \overline{X}, \alpha\}$ iff α modifies $\overline{X}$.

This bijection limits modifiers to sisters of $\overline{X}$s. It is redundant with the Projection Principle, which blocks modifiers from complement position. But it also blocks modifiers from being in Specifier positions, something that would otherwise be available. It also forces sisters to $\overline{X}$s to be modifiers, and this is the use we have of it here. Because expletives, by virtue of being semantically vacuous, cannot modify, they will be banned from this position.

There are other ways of blocking (11b), of course, and there is no particular reason to believe that this is the correct method. But let's adopt this principle until something better comes along.

There is one last fact in this domain that requires addressing. This is that when there is no subject θ-role, only an expletive can satisfy the Extended Projection Principle. Placing an argument in the Specifier of IP in such a case is ungrammatical, as a comparison between (7) and (13) indicates.

(7) It seems that we are behind.

(13) * Jerry seems that we are behind.

Many formulations of the Theta Criterion are designed to capture this fact, requiring that there be a θ-marked position for each argument. We have a different option. Because modification is restricted to just adjunct positions, semantically contentful phrases in specifier positions are not going to be able to modify. If the only other way a meaningful phrase can be put together with the rest of a sentence is by way of a θ-role, then the ungrammaticality of (13) will follow. Under either way of thinking, the point is the same. If NPs in Specifier of IP can combine semantically with the rest of the material in the sentence only by way of a θ-role, then that NP will have to have no semantic content (be an expletive) or get a θ-role.

Let's return now to our original set of problems: (a) how do subjects get their θ-roles, (b) when there are two or more complements, what determines their order and (c) what are we to do about the paradoxical position of CP objects?

3.2 Movement

At present our system imposes a linear order on complements and non-complements. Because complements are trapped inside the lowest $\overline{\text{X}}$ and non-complements are forced out of that $\overline{\text{X}}$, complements will precede non-complements when they both follow the head. More generally, (14) holds.

(14) The Projection Principle entails: If an argument, X, and a non-argument, Y, both fall linearly on the same side of the head, then X will come closer to the head than Y.

It's this generalization that the problem with respect to CP objects is tangled up in. We find that object CPs must be part of the smallest $\overline{\text{V}}$ they are an object of — this is what is responsible for the inability of object CPs to be outside the $\overline{\text{V}}$ that *do so* occupies in examples like (15) — and yet they tend to follow modifiers that are able to stand outside this $\overline{\text{V}}$, as (16) indicates.

(15) * Max said that pigs fly and Sally did so that snakes swim.

(16) a. Max said yesterday that pigs fly.
b. * Max [said that pigs fly] yesterday.

The evidence concerning the hierarchical position of CP objects is at odds with what (14) says about their linear order.

In fact, as briefly noted in the previous chapter, this problem arises with other kinds of objects as well. Object NPs also seem to be forced to reside in the lowest $\overline{\text{V}}$, as indicated by their inability to be stranded by *do so* anaphora as in (17).

(17) * Max ate the rotting kumquats, but Sally hasn't done so the delicious strawberries.

And, as expected, it does seem that (14) holds for some kinds of NP objects, as in (18) and (19).

(18) a. Jill ate it at noon.
b. * Jill ate at noon it.

(19) a. Jill ate squash at noon.
b. * Jill ate at noon squash.

But in other situations, it's less clear, or downright wrong, as in (20) and (21).

(20) a. Jill ate the rotting kumquats at noon.
b. Jill ate at noon the rotting kumquats.

(21) a. ?? Jill said [that you shouldn't eat kumquats] at noon.
b. Jill said at noon [that you shouldn't eat kumquats].

And a similar situation arises in cases where the object is a PP as well. The ungrammaticality of (22) indicates that object PPs are trapped within the lowest $\overline{\text{V}}$ they are the object for, and yet as (23) indicates, they can follow modifiers of this $\overline{\text{V}}$.

(22) * Max put it on the table and Sally has done so on the chair.

(23) a. Max put it on the table carefully.
b. Max put it carefully on the table.

There are other situations in which violations of (14) arise. One that has engendered a lot attention are constituent questions of the type that (24) illustrates.

(24) Which linguist has given up?

One feature of such questions is that they have an NP, or other phrase, that contains a special interrogative word. In (24), that interrogative word is the determiner *which*. Another feature of such questions is that a phrase containing that interrogative word — known as the "interrogative phrase" or "wh-phrase" — appears at the very beginning of the sentence that constitutes the question. In (24), this second requirement is trivially satisfied as the wh-phrase is appears at the beginning of the sentence by virtue of its status as subject. In other cases, however, this requirement results in a violation of the expected position of the wh-phrase. In (25), for instance, the wh-phrases are objects and yet appear at the beginning of the sentence.

(25) a. Which linguist has Max visited?
 b. On which table has Max put it?

Although the linear position of these objects is not where we would expect them to be, they still behave as if they are within the smallest $\overline{\mathrm{V}}$; the *do so* anaphora test behaves in just the same way it always does for these phrases.

(26) a. * I know that Max visited some philosopher, but I don't know which linguist he has done so.
 b. * I know that Max has put it on some table, but I don't know which table Sally has done so.

Here too, then, we have a violation of (14).

There have been a number of approaches to these problems. I will not be able to consider them all in these lectures. The first we will consider is the first that was proposed. Zellig Harris, and then Noam Chomsky in many articles in the late 1950's and early 1960's, argued that the solution to these violations of (14) involves giving sentences more than one phrase-marker. In addition to the D-structures we have arrived at, there are (sometimes) other phrase markers that are formed from D-structures by way of "Transformations." In the cases at hand, the Transformations are "movement rules," that preserve everything in the D-structure of the sentence affected except for the position of one phrase. The phrase-markers that are formed by these Transformations are the ones that get pronounced, and so the position that they give to the phrases they move are the ones that determine their linear position. This phrase marker is called a "Surface Structure," or S-structure.

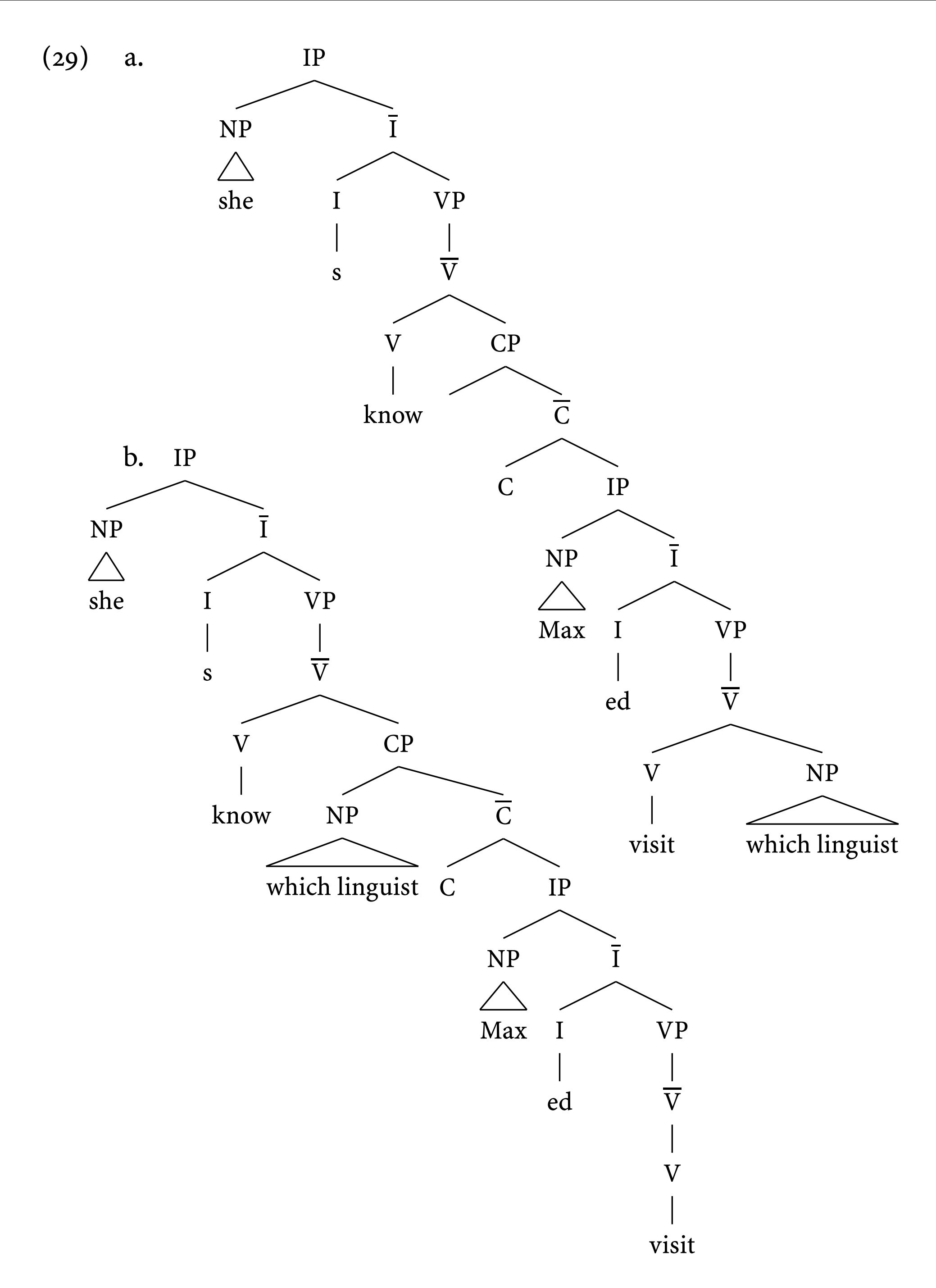
In the case of constituent questions, the movement rule involved is known as "Wh Movement," and a preliminary formulation of it is in (27).

(27) Wh Movement
Move a wh-phrase into Specifier of CP.

We will see evidence in the following chapter that the position that the wh-phrase is moved to is the Specifier of CP. At this stage we can note that this decision correctly places the wh-phrase at the very beginning of the question. We might also note that this provides an answer to one of the problems we left open at the end of the previous chapter: what is the material that occupies Specifier of CP. It appears that the material that fits into the Specifier of CP is always material that has been moved from elsewhere. We will see other cases of movement to Specifier of CP, by Wh Movement seems to be one that is found in many languages beyond English. This rule would give the D-structure/S-structure mapping in (29) on the next page for a sentence like (28).

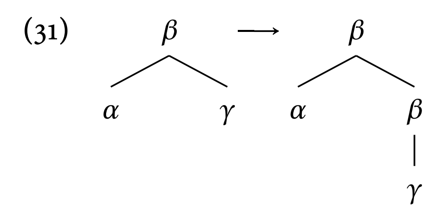
(28) She knows which linguist Max visited.

For the cases that pose a problem for (14), Ross (1967) made an influential proposal. He proposed a rule sometimes called "Heavy NP Shift," which forms from sentences where the argument precedes the non-argument, sentences where that argument has been "shifted" to the right. His rule can be formulated as (30).



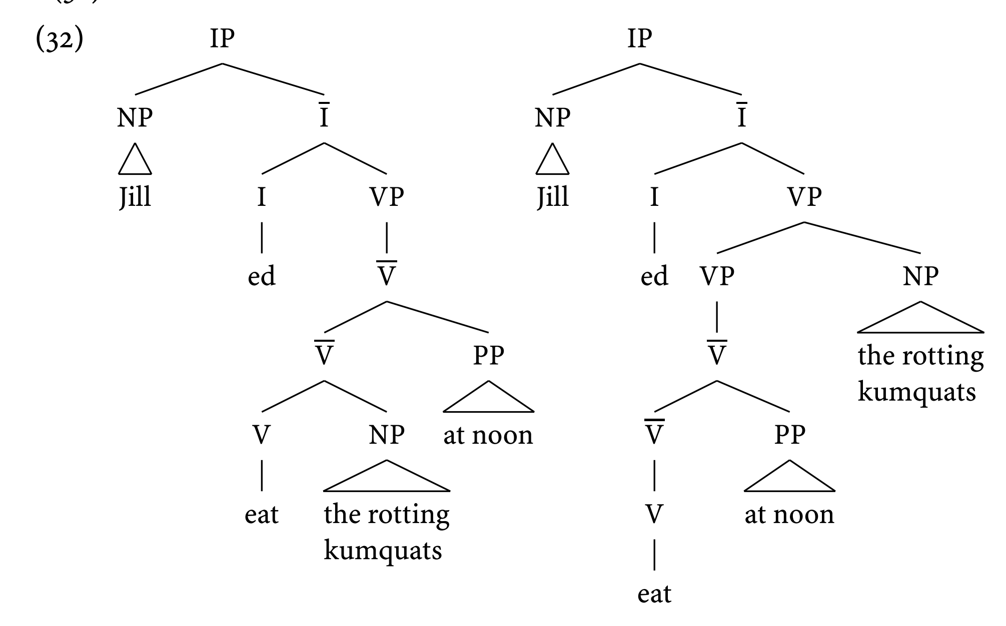
(30) Heavy NP Shift
Right adjoin a phrase to the first VP containing that phrase.

By "Adjunction" is meant a process that takes two terms and makes them both daughters of a newly created phrase which is identical to and replaces one of the two terms brought together. By convention, the term replaces is the object of the expression "x is adjoined to y." In diagram form, "α adjoins to β" is (31).

(31)


To "right adjoin" α to β is to cause the linearization algorithm to force α to the right of β. To "left adjoin" α to β is to cause the linearization algorithm to place α to the left of β. When the direction of adjunction is left unspecified, it is understood to be leftward, as in (31).

Heavy NP Shift would form an S-structure from the problematic (20b) from a D-structure that, if pronounced, would correspond to (20a). This is sketched in (32).

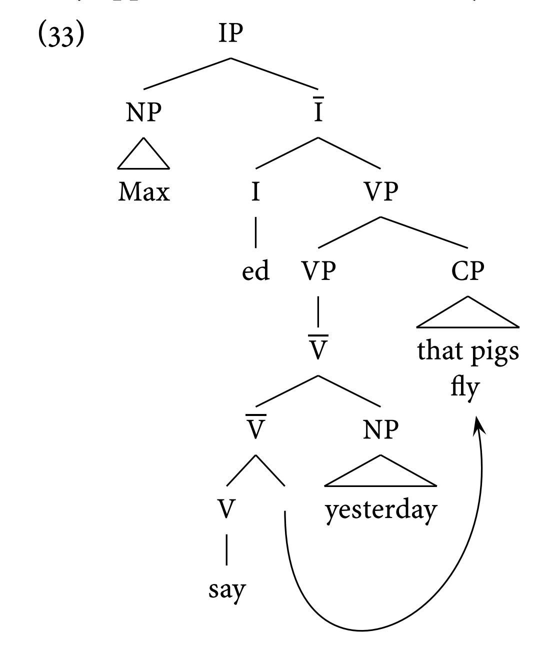
(32)


This sentence, then, has two parses. One that satisfies the Projection Principle and the Theta Criterion, and the other which doesn't, but is mapped from the first by Heavy NP Shift. This sentence is a series of phrase markers, then, made up of the two parses shown above. We can see the members of this series as being generated in the following fashion: the first member is produced by group

and satisfies the Projection Principle, the Theta Criterion, the conditions on modifier type and the $\overline{\text{X}}$ Skeleton. This is the "D-structure." The other parses in the series are produced by the action of transformational rules acting on elements already found in the series. In (32), there is only one other such phrase marker, but we will see cases soon in which transformational rules apply more than once forming series of greater length. We will eventually have to control the method by which transformations produce additional elements, but in the simple cases we will be concerned with right now, there are only two members of the series: the D-structure and another formed by a transformation acting on that D-structure. We speak the final phrase marker in the series. In the case of NP Shift, this is the parse produced by the transformation. The parse that is spoken is called the "S-structure," and the series of parses are called "Derivations."

Heavy NP Shift will have to include conditions that generate the paradigm of contrasts that we encountered earlier. We will want to ensure that it obligatorily applies in cases where the object is a CP, as in (33).

(33)



(I will frequently collapse two adjacent phrase markers in a derivation into one parse, using the device of a movement arrow, as in this example.) By contrast, we must find a way of blocking Heavy NP Shift from applying when the object is an NP and "short": a pronoun or bare plural, for instance. For other objects, Heavy NP Shift seems to be an option, with there being preferences determined

by the length of the object relative to the material around it. I don't know the source for these preferences. I speculate that the conditions governing these preferences come from something other than syntax proper.[1] The conditions on Heavy NP Shift, then, is as indicated in (34).

(34) a. Obligatory Heavy NP Shift:
  i. (finite) CPs:
    * Max [said that pigs fly] yesterday.
    Max said yesterday that pigs fly.
 b. Optional Heavy NP Shift:
  i. "full" NPs:
    Max visited the students yesterday.
    ? Max visited yesterday the students.
  ii. PPs:
    Max talked to the students yesterday.
    Max talked yesterday to the students
 c. Blocked Heavy NP Shift:
  i. Pronouns:
    Max visited him yesterday.
    * Max visited yesterday him.
  ii. "short" NPs:
    Max visited children yesterday.
    ?* Max visited yesterday children.

This provides a characterization, then, of the cases in which a phrase seems to be hierarchically in a position different from where its linear position would lead us to expect. This movement transformation hypothesis claims that there is no real deviation from the expected relationship between hierarchical position and linear position in these cases. What makes them exceptional is the existence of a phrase marker at odds with the conditions we have discovered on D-structures. This other phrase marker gives certain phrases a non-D-structure hierarchical position, and it's this new position that reflects the linear position of the affected phrase.

1 Some plausible candidates are conditions on processing load — the longer an object gets the more a configuration triggering the effects of minimal attachment discussed in the first chapter arise — conditions on the prosodic interpretation of syntactic structures, and conditions on how the distinction between "old" information and "new" information is signaled.

This is a deeply mysterious affair. Why should movement transformations exist? We should hope to find reasons for the existence of these deviations from the normal relationship between hierarchical position and linear order. If the movement transformation approach to these cases is correct, then the mystery becomes the question why certain phrases should be given two positions within a phrase marker. A goal of much current research in syntax is to find an answer to this mystery. We will have a chance to look at some recent attempts in what follows.

3.3 The Underlying Order of Complements

Let's turn now to the question of what fixes the order of complements in D-structure. We'll focus here on the question of what determines the linear position of the objects of verbs. Recall that the way in which complements to verbs are ordered does not vary with the choice of verb. We will not want to encode the order of complements on a verb by verb basis, then. Instead, the relevant factor in determining the order of objects seems to be the categorial type that the objects are. In general, as our initial phrase structure rules encoded, verbal complements are ordered as follows.[2]

(35) NP + PP + CP

(36) a. I told [$_{NP}$ her] [$_{CP}$ that she should join].
b. * I told [$_{CP}$ that she should join] [$_{NP}$ she].
c. Mary explained [$_{PP}$ to me] [$_{CP}$ that I should join].
d. ?* Mary explained [$_{PP}$ that I should join] [$_{PP}$ to me].
e. Sally told [$_{NP}$ it] [$_{PP}$ to Bill].
f. ?? Sally told [$_{PP}$ to Bill] [$_{NP}$ it].

We already have an ingredient in the solution to this problem. Because Heavy NP Shift is obligatory for (finite) CPs, it will force these CPs to follow all other complements. The ungrammaticality of (36b) and (36d) can be traced back to the obligatoriness of Heavy NP Shift of the CP in these examples. What we are left with, then, is finding a reason that NP objects must precede PP objects.

In his 1981 dissertation, Tim Stowell argued for an answer to this question that has been influential. He suggested that systems of "Case assignment" are

2 As a simplifying measure, we consider only finite clauses.

responsible for ordering these terms. NPs differ from other kinds of phrases in Indo-European in being able to host Case morphology. In English this happens with pronouns only: the pronouns *she* and *her*, for instance, have the same meaning but have a different "Case" form. The particular Case borne by a pronoun is determined by its syntactic position. In languages that are richer in Cases than English is we can see that the Case borne by a NP is determined by a term proximate to the NP. In German, for instance, a certain class of prepositions and verbs determine Accusative Case for their complements, while others determine Dative Case. It is also often the case in other Indo-European languages that NPs other than just pronouns can bear Case morphology — sometimes on the head of the NP, sometimes on the determiner, sometimes spread across modifiers of various sorts within the NPs. Let's imagine then that, in general, NPs must be related to Case assigners. Or:

(37) Case Filter
An NP must be assigned Case if it is an argument.

I've restricted the Case Filter to argument NPs because, as we'll see, adjunct NPs do not seem to be positioned in a way that suggests they are sensitive to Case assignment. Moreover, typically the Case morphology they bear is fixed, and not sensitive to Case assigners. The Case filter, then, requires that some parse in a sentence's derivation puts every argument NP that sentence contains in a Case marked position.

As I noted before, Case marked positions are ones that are close to terms that are responsible for assigning (i.e, determining) the Case. So, all we have to do now is know what those terms are and what "close" means, and we'll be able to use the Case filter to distinguish sentences in terms of grammaticality. The "object" Cases — so-called Accusative and Dative and Locative, and a host of others — are assigned by particular lexical items. In English there are just three cases: accusative, nominative and genitive. The form of pronouns in English depends on this Case information, as well as the number, person, and, in some cases, gender of the referent of the pronoun.

	Nominative	*Accusative*	*Genitive*
1st, sing	I	me	my
1st, plur	we	us	our
2nd, sing	you	you	your
2nd, plur	you	you	your
3rd, sing, fem	she	her	her
3rd, sing, masc	he	him	his
3rd, sing, neut	it	it	its
3rd, plur	they	them	their

Our immediate task, then, is to determine what terms are responsible for assigning these Cases. The genitive appears on NPs that are in the Specifier of NP in English, as in (38).

(38)
a. **your** hat *compare:* ***your** left.
b. **his** marbles *compare:* *I talked to **his**.
c. **our** problems *compare:* *She met **our**.

We put off for a couple chapters investigating what the item is that assigns Genitive; there are facts about the shape of NPs that we have yet to discover that are relevant.

Nominative Case is found on the "subjects" of finite IPs, as in (39).

(39)
a. **She** left. *compare:* *I left **she**.
b. **We** arrived before **he** left. *compare:* *I talked to **he**.
c. **They** thought **I** would stay. *compare:* ***they** hat

Finite IPs are distinguished from other phrases by the presence of a finite I^0: tense/agree inflection or a modal. Moreover, the nominative Case-marked NP appears in Specifier of IP, at least in the examples we have examined so far. (We will encounter soon cases that don't meet this description.) An initial stab at the conditions on nominative Case assignment, then, is (40).

(40) Finite I^0 assigns Nominative Case to its Specifier position.

Accusative Case is found on NPs that immediately follow some verbs and prepositions, as in (41).

(41)
a. We met **him**. *compare:* ***Him** left.
b. They talked to **me** *compare:* *They are happy **me**.
c. She stood near **us**. *compare:* ***us** hats.

It appears that verbs and prepositions are the terms that assign accusative Case. Tim Stowell argued that one of the conditions on accusative Case assignment is adjacency. We will see soon that additional conditions are required too. An initial stab at the conditions on accusative Case assignment, then, is (42).

(42) Accusative Case is assigned to the position adjacent to a verb or preposition.

Stowell (1981) proposes to derive the ordering of complements by way of the Case Filter and an adjacency condition on accusative Case assignment. Because NPs are subject to the Case filter, when they get a θ-role from a verb, they will necessarily be positioned adjacent to this verb. By contrast, PPs are not subject to the Case filter and they are free to appear anywhere in the $\overline{\text{V}}$ that contains their θ-role assigner. This will ensure that when an NP and PP share a $\overline{\text{V}}$, the NP will come closer to the verb than does the PP. In a language like English, in which the linearization parameters cause the verb to be first in its $\overline{\text{V}}$, this will mean that an object NP will precede an object PP.

With regard to the relative order of PP and finite CP, Stowell uses Ross's Heavy NP Shift operation, but he suggests that Case plays a role in making Heavy NP Shift obligatory. He speculates that finite CPs, like NPs, must receive Case but that unlike NPs they cannot sit in Case-marked positions at S-structure. He calls this the "Case Resistance Principle."

(43) Case Resistance Principle
A (finite) CP may not be in a Case marked position at S-structure.

The Case Resistance Principle requires finite CPs to be moved by S-structure to some non-Case marked position. NP Shift is capable of doing this, and thus, in cases where a complement CP shows up string finally in a VP, it has satisfied Stowell's injunction against surfacing in Case marked positions by undergoing NP Shift. There is, surprisingly, a certain amount of evidence for this picture.

Consider, for example, situations where the finite clause is the "subject" of the clause. Here too, following an argument from Koster (1978), we see that there is some reason for thinking that it isn't actually in the Nominative Case-marked, Specifier of IP position. Koster's argument makes reference to a process that is found in certain question-types in English. Normally, in these question contexts, it is possible to move I^0 to the front of a sentence, as in (44).

(44) Mary will put the book on the table. →
Will Mary put the book on the table?

How precisely this is done is the subject of the next chapter. What's relevant here is that this process is blocked if it interacts with another process that moves something to the left edge of IP, as in (45).

(45) Mary will put the book on the table. →
On the table, Mary will put the book.

These two processes cannot apply to the same sentence, as (46) indicates.

(46) a. Mary will put the book on the table →
b. On the table, Mary will put the book. →
c. * Will on the table, Mary put the book?

Now, interestingly, for many speakers of English the presence of a finite CP as a subject of a sentence also blocks movement of I^0.

(47) a. That Mary has left should be obvious. →
b. * Should that Mary has left be obvious?
c. That Sally sleeps late bothers Mittie. →
d. * Does that Sally sleeps late bother Mittie?

This would follow if finite CPs are driven from the nominative Case marked Specifier of IP, and adjoin to the left of IP in these cases. Stowell's method of forcing NP Shift to apply to complement CPs would extend to this scenario as well. CPs start out in the nominative Case-marked position, but are driven from there in order to form an S-structure.

This evidence all points to the fact that finite CPs move to either the left or right linear edges of the sentences they are part of. If CPs are prevented from being in Case marked positions in the S-structure, this would get us close to deriving this fact. But is there evidence for the motivation for this movement that Stowell proposes? In particular, is there motivation for the claim that finite CPs, like NPs, require Case?

One piece of suggestive evidence comes from the class of verbs that permit both NPs and finite CPs. These are only just those verbs that already exceptionally allow two NPs. These include verbs like *promise*, *tell*, and *show*; they are known as "double object" verbs. Some examples indicating that double object verbs also take NP and CP are in (48).

(48) a. Mary promised me that she would sing.
Mary promised me the ring

b. Jerry told me that he can't stand Mary's singing.
Jerry told me the story.

c. Sheila showed me that she cares.
Sheila showed me her concern.

This isn't completely the case, as (49) is a counterexample.

(49) a. Mary persuaded Bill that he should go.

b. * Mary persuaded Bill the fact

But, so far as I know, (49) is the only counterexample. To the extent that there is a match in the verbs which accept NP CP and those which accept NP NP complements, there are grounds for believing that their surface positions are governed by the same, or similar, principles. And to the extent that the dominant principle is the Case Filter, then there is reason to conclude from these data that CPs are subject to the Case Filter as well.

This match between the distribution of NPs and CPs should be completely general if Stowell is correct. Indeed, finite CPs are distributed in sentences much like NPs are. We find them as complements to verbs, as we have seen, and in the subject position of other finite clauses, but not in the subject position of infinitives (as we shall see shortly). These are just the Case-marked positions. But there are several differences in their distribution. In English, finite CPs are probably never found as the complements to a preposition, though of course NPs are. The only potential counterexample comes from temporal prepositions, as in (50).

(50) a. I left before Mary arrived.

b. * I left before that Mary arrived.

Similarly, both adjectives and nouns can take CP complements, but not NP complements.

(51) a. Sean is unhappy that he had to sing.

b. * Sean is unhappy that.

(52) a. the proof that lemons cure cancer

b. * The proof this fact.

If this has a Case Theoretic explanation, then Stowell's theory is in trouble. But it could also be that this arises because of some property of c-selection. The jury is still out with respect to Stowell's suggestion that Case is responsible for the surface position of CPs.

Though the evidence suggests that Stowell's ideas meet with some success, there are problems too. One which threatens Stowell's Adjacency Condition on Case assignment, and its use in fixing the order of complements, concerns so-called "double object" constructions, as in:

(53) a. Mary showed Bill the picture.
  b. Bill baked Sally a cake.

How can the second NP in these examples receive Case? We will have to develop some of the rest of the system that is responsible for linearizing arguments before we can engage this difficulty. So, let me ask for your indulgence and postpone examining this case.

3.4 Small Clauses and the Derived Subjects Hypothesis

Of the problems we set out to answer at the beginning of this chapter, we've now engaged the majority of them. The paradoxical behavior of object CPs we've given an account of that involves movement transformations. The fact that NP objects precede object PPs, modulo the action of Heavy NP Shift, we've credited to the Case filter in combination with an adjacency condition on accusative Case assignment. We're left with the question of what resides in the Specifier positions of certain phrases. We've seen that Specifier of CP is occupied by a wh-phrase, moved to that position by Wh-movement. We've also found the contents of Specifier of IP and NP: these are the subjects and determiners (or genitive Case-marked NPs), respectively. We're left with the following phrases.

(54) What resides in the Specifiers of:
  a. Adjective Phrase
  b. Verb Phrase
  c. Prepositional Phrase
  d. Determiner Phrase
  e. Adverb Phrase

We will return to question concerning Determiner Phrases in a later chapter. The structure of Adverb Phrases is not well-studied, partly because of the apparent poverty of material that can be in adverb phrases. We will generally ignore the structure of these phrases as a result. It's time, though, to see what occupies the Specifier positions of APs, VPs and PPs.

Let's begin with VPs and APs. What is it that distinguishes VPs from $\overline{\text{V}}$s and APs from $\overline{\text{A}}$s? One hypothesis is that VPs and APs can have "subjects" in them, and $\overline{\text{V}}$s and $\overline{\text{A}}$s can't. This can be seen in certain special circumstances, of which (55) are examples.

(55) a. She will let [them make a cabinet].
b. This will make [them angry at me].

As indicated by the brackets, it is thought that the strings following *let* and *make* form a single phrase which is serving as the complement to these *let* and *make*.

Why don't we instead believe that these verbs are followed by two complements, the first of which is *them*? Because *them* is not an argument of the verb. We can see from other examples that this isn't the case. In (56), for example, this position can be occupied by an expletive.

(56) a. I will let [it seem that there is a mistake on the handout].
b. This will make [it obvious that there is a mistake on the handout].

We discovered in 3.1 that expletive *it* can only stand in Specifier positions. Thus, the positions following *let* and *make* in these cases must be Specifier positions.

Note that the bracketed string in (55) is, indeed, an AP and VP, and not IP. This, we know, because the strings which follows these specifier positions are characterized by the rules which generate $\overline{\text{V}}$s and $\overline{\text{A}}$s, and not those which characterize $\overline{\text{I}}$s. As (57) indicates, an I^0 is not found in this string.

(57) a. * I will let [her should make a cabinet].
b. * I will let [her to make a cabinet].
c. * I will let [her makes a cabinet].

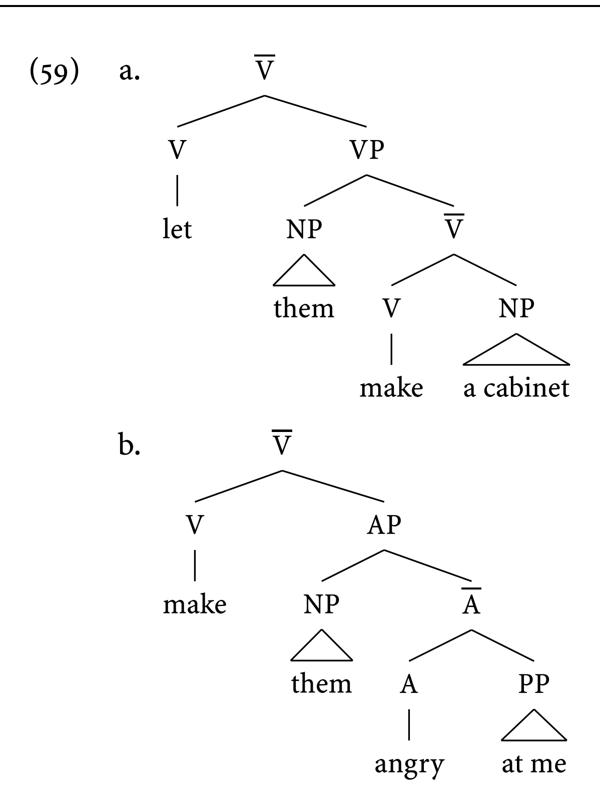
Finally, why might wonder if the bracketed strings in (55) are just $\overline{\text{A}}$ and $\overline{\text{V}}$. Why don't we think that *them* is inside the $\overline{\text{V}}$ or $\overline{\text{A}}$? I can't show you a reason for this in the case of $\overline{\text{A}}$; but for $\overline{\text{V}}$ we can tell this from the action of the *do so* phenomenon. Recall that *do so* replaces $\overline{\text{V}}$s; so the contrast in (58) indicates that the NP in (55a) is not in $\overline{\text{V}}$.

(58) a. I will let [her make a cabinet] and I will let [him do so] as well.
b. * I will let [her make a cabinet] and I will let [do so] as well.

From these observations we will want to give the examples in (55) the parses in (59).

(59) a.

b.



These complements to *let* and *make* are sometimes called "small clauses," and it is to Stowell, once again, that we owe this way of viewing them.[3] If this is correct, we have now discovered what stands in the Specifier of AP and VP. These positions can contain a "subject" argument.

A similar conclusion can be reached for PPs, though here the evidence is somewhat less meager. Cases such as (60) plausibly involve small clause PPs.

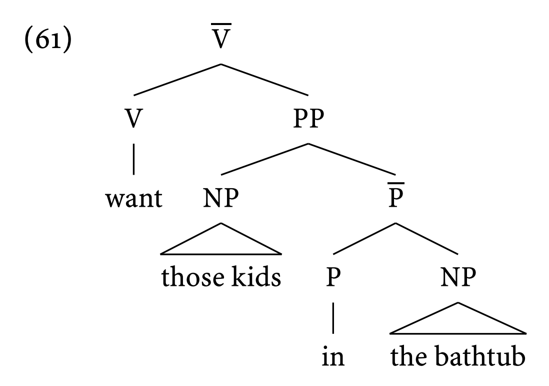
(60) a. I want those kids in the bathtub!

b. She saw them in the next room.

Both of these examples can have a meaning in which it seems that the verb has just one object. In (60a), the object of *want* seems to be, roughly, what *them in the bathtub now* describes. The content of my desire is for those damn kids to be in the tub. Under our present understanding of how θ-roles are assigned, this requires that (60a) have the parse in (61).

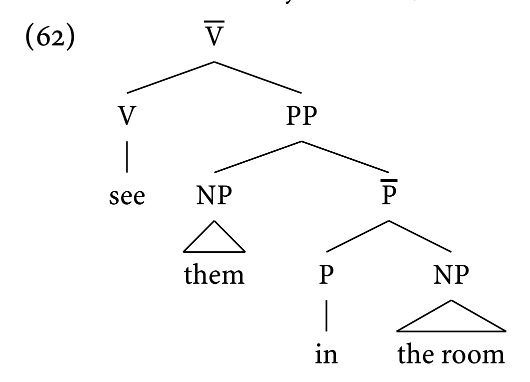
3 See Stowell (1978).

(61)



Similarly, among its several meanings, (60b) can describe a situation in which what was seen was *them in the room*. This reading arises when *in the room* does not describe where the seeing event took place — that reading is manufactured by letting *in the room* modify *see* — but instead describes the location of the referent of *them*. To produce such a reading, we want *them in the room* to be the object of *see*, and this is achieved by a parse like (62).

(62)



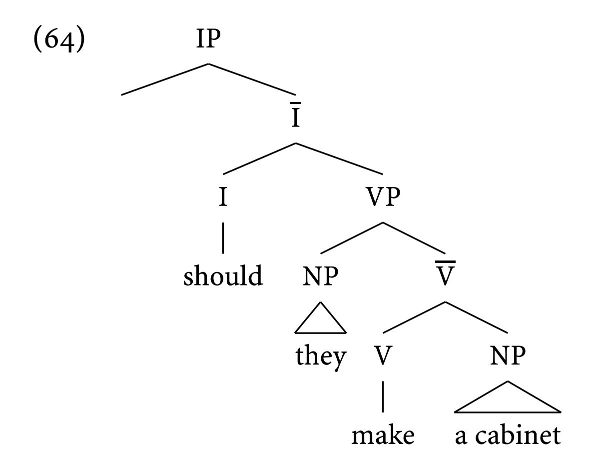
In both of the examples in (60), then, there is reason to believe that the verbs are followed by one PP, whose Specifier is filled with the NP argument.

Consider now the examples in (63).

(63) a. They should make a cabinet.
 b. They are angry at me.
 c. They are in the next room.

These sentences contain the same VP, AP and PP that constituted small clauses in our earlier examples. Moreover, the θ-role borne by the subject in (63) seems to be the same as that borne by the NP in the examples with small clauses. Indeed, the meaning of the sentences in (63) seems to be identical to those of the small clauses in the previous examples, except for that meaning introduced by

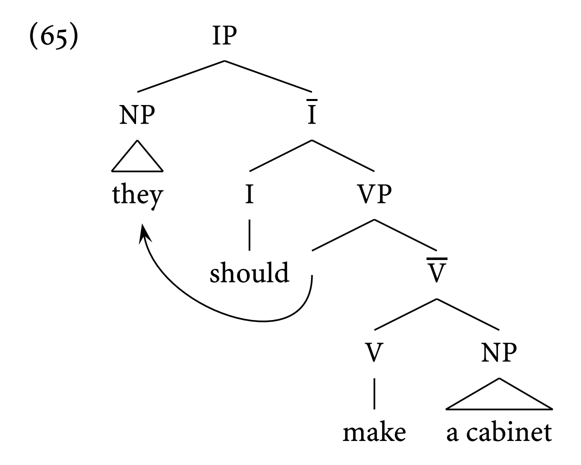
the modal *should* (in (63a)) and the verb *be* (in (63b,c)). A number of linguists independently noted that there could well be a small clause within (63), but that this arrangement is obscured by a movement transformation that relocates the subject.[4] Indeed, the Case filter based theory about where argument NPs must be predicts that (63a) could have a small clause in it at D-structure. Suppose that the θ-role that *they* bears in (63a) is assigned to the same position that it is in our small clause example in (55a), that is to the Specifier of VP. Then the Theta Criterion would require the D-structure of (63a) to be (64).

(64)


The subject argument, *they*, bears nominative Case, and yet in this phrase marker, it is not in the position our system assigns nominative Case to. That position is the Specifier of a finite IP, which is the position that *they* is, in fact, spoken in. Imagine, then, that there is a movement transformation that relocates *they* into the Specifier of IP, as in (65).

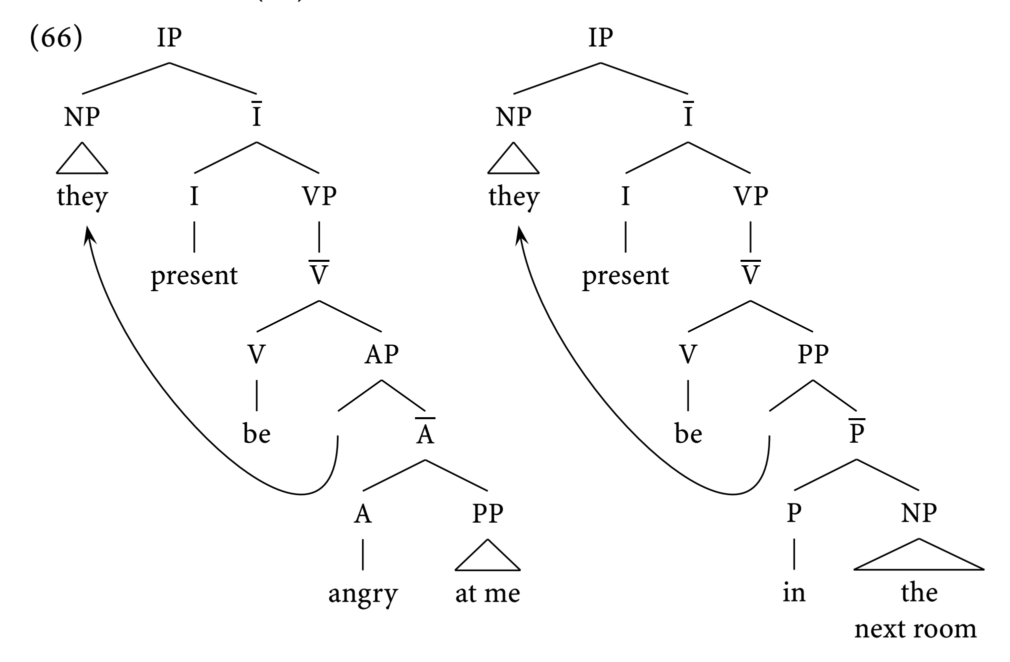
4 See Kuroda (1988), Kitagawa (1986), Fukui and Speas (1986) and Koopman and Sportiche (1991) for proposals along these lines that fit into the general theoretical assumptions defended here, and Fillmore (1968) and especially McCawley (1970) for early antecedents to the idea.

(65)



The same derivations could be present in (63b) and (63c) as well. They would be as indicated in (66).

(66)



The conjecture that the subject arguments of IPs start out inside small clauses and move into the position they are spoken goes by the name of "the derived subjects hypothesis." The movement transformation responsible for relocating the subjects is known as "Argument Movement," or more commonly: A-Movement.

(67) Argument Movement
Move an XP into Specifier position.

We shall examine some evidence in support of the derived subjects hypothesis in section 3.6 after we see other environments in which A Movement is alleged to occur. But before we turn to those tasks, we need to sharpen the conditions on Case assignment in order for them to produce the desired effects under the derived subjects hypothesis. Our present conditions on Case assignment are (68).

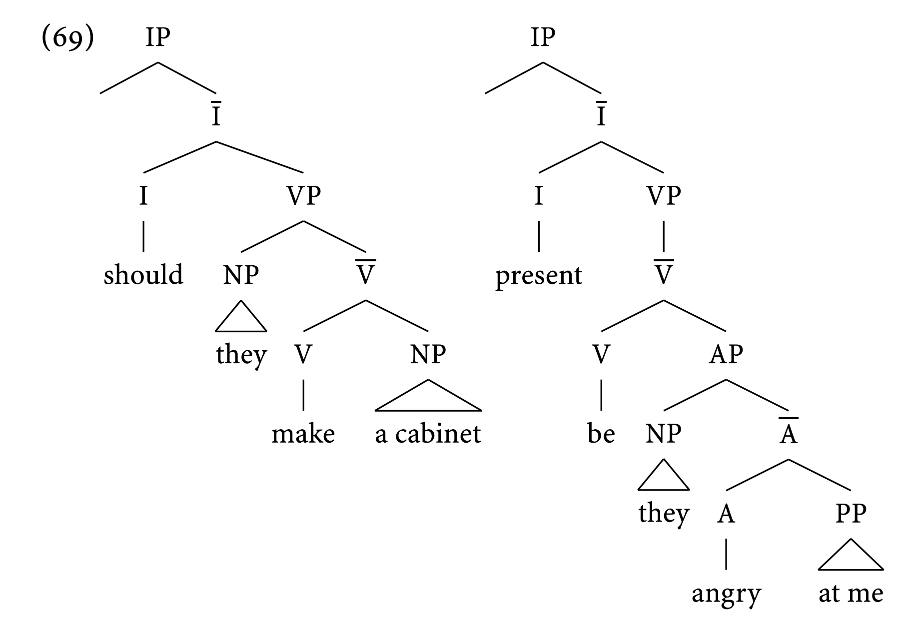
(68) Conditions on Case assignment

a. Nominative Case is assigned by finite I^0 to its Specifier position.

b. Accusative Case is assigned by verbs and prepositions to an adjacent position.

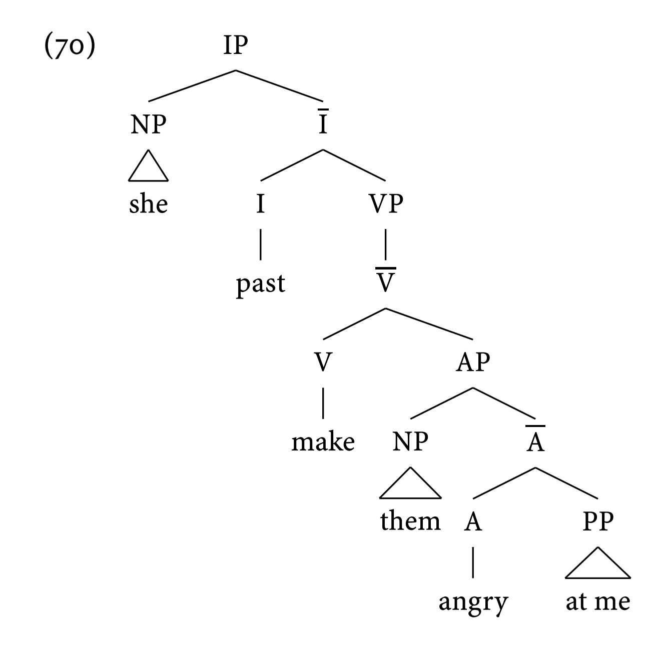
The description of where accusative Case is assigned does not produce the right results when coupled with the derived subjects hypothesis. To see this, consider the D-structures of some of the examples we've looked at.

(69)



The subjects in these examples cannot be assigned accusative Case; they only appear with nominative. But the statement of accusative Case assignment in (68b) does not ensure this. The subject is adjacent to the verb *make* in (69a), and in (69b) it is adjacent to the verb *be*. The case in (69b) contrasts minimally with (70).

(70)



In (70), *make* assigns accusative Case to *them*. Because of the parallel structures of (70) and (69b), the usual conclusion is that the difference in availability of accusative Case must devolve to the difference in the verbs involved. Unlike *make*, *be* does not assign accusative Case. It appears that (71) is true.

(71) The ability to assign accusative Case is part of the lexical specification of a verb.

The response to (69a) needs to be more general. No matter what verb is involved, it is not able to assign accusative Case to its Specifier position. The usual answer given to this problem, one that extends to other cases, as we shall see, is based on the structural relationship between the verb and the positions it can assign Case to. A verb is able to assign accusative Case to its sister, as in (69a), or to a position within its sister, as in (70). To capture this range of positions, we define the following relation.[5]

5 This relation comes from Reinhart (1976), whose definition is slightly different from the one produced here. (The central difference is that she uses "branching node" in place of my "phrase.") The name "c-command" is short for "constituent command." Reinhart presents her proposal as a modification of the relation "command" in Langacker (1969). It is close to the notion "in construction with" that is found in Klima (1964), which Langacker cites as his model.

(72) α c-commands β iff
  a. every phrase dominating α also dominates β, and
  b. α does not dominate β.

The c-command relation governs a wide range of syntactic phenomena, and so it will arise in several other contexts in these lectures. The definition in (72) anticipates these later uses, and so goes a bit beyond what is necessary for the Case assignment relation. Only (72a) is needed here. (72b) prevents things from c-commanding themselves or their contents, and neither of these scenarios are relevant for the Case assignment relation. Armed with c-command, we can restrict the condition on accusative Case assignment to (73) and, thereby, prevent verbs from assigning accusative Case to their Specifier position.

(73) Accusative Case is assigned by some verbs and prepositions to positions that they c-command and are adjacent to.

We need one additional change to the condition on accusative Case assignment. This change is made necessary by the contrast in grammaticality of examples like (74).

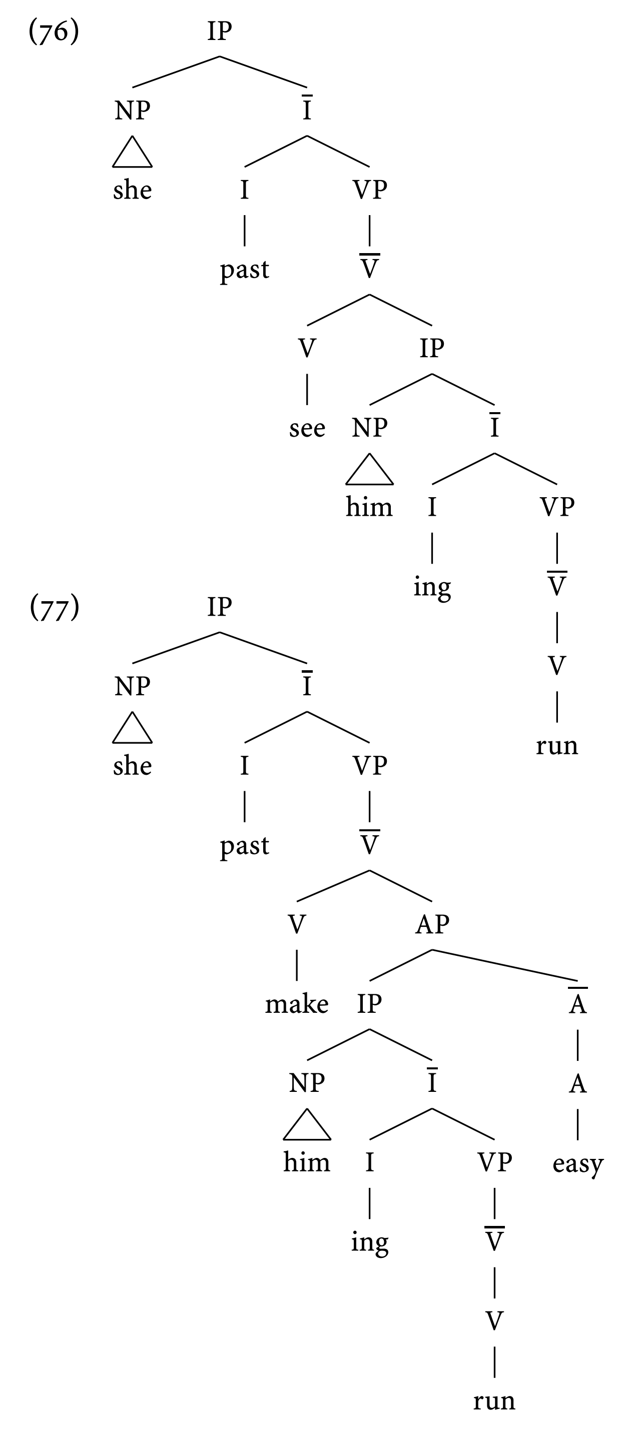
(74) a. * She made him running easy.
  b. She saw him running.

The complement to *saw* in (74b) is a kind of non-finite IP called a "gerund." These IPs are headed by the inflectional morpheme *ing* that appears on the following verb. The subject of this IP in (74b) receives its accusative Case from *see*; its S-structure parse is (76) on the facing page. The condition on accusative Case assignment in (73) correctly captures the grammaticality of this example; *see* c-commands and is adjacent to *him*.

Consider now the contrasting (74a). In this example, the gerund is the subject of a small clause, as indicated in (77) on the next page. The condition on accusative Case assignment in (73) would allow *make* to assign Case to *him* in (77): *make* c-commands, and is adjacent to, *him*. And yet, the ungrammaticality of this example traces back to the availability of Case for *him*, as can be appreciated by considering the fact that this example becomes grammatical without *him*.

(75) She made running easy.

The ungrammaticality of (74a) should arise because it violates the Case filter, and this indicates that *make* cannot assign accusative to *him*.



A salient difference between (74a) and (74b) is the depth of embedding of the NP to be assigned Case. There appears to be a condition on how far into

its sister a verb can assign Case. The precise statement of this condition is still under investigation; we will consider some of the proposals in what follows. I will add this condition to the formulation of Case assignment in a way that allows us to modify it as we go along. An initial reformulation, then, is (78).

(78) Accusative Case is assigned by some verbs and prepositions to positions that they c-command, are klose to, and are adjacent to.

α is KLOSE to β iff there is no more than one phrase that dominates β but not α.

This definition allows a verb to assign Case to its sister and to the immediate daughters of its sister, but prevents it from assigning Case to any position more deeply in its sister. In *Lectures on Government and Binding*, and much subsequent work, a term — "government" — is defined that collapses the c-command and klose relations. The condition on accusative Case assignment is often formulated as (79).

(79) Accusative Case is assigned by some verbs and prepositions to positions that they govern and are adjacent to.

α GOVERNS β iff α c-commands β and α is klose to β.

With these changes, then, the system of Case assignment and the Case filter correctly work with the derived subjects hypothesis to ensure that subjects of finite clauses show up with nominative Case in the Specifier of IP at S-structure. The existence of small clauses, and the accompanying thesis that subjects of IPs start out in small clauses, requires that the subject θ-role be assigned to the Specifier of the phrase headed by the term that has the θ-role. The Projection Principle determines where θ-roles are assigned, and so we should modify this condition to capture where subject θ-roles are assigned. At present, the Projection Principle requires that words assign their object θ-roles to positions that they are sisters to. Let's broaden this to:

(80) The Projection Principle
α assigns a θ-role to β iff γ and β are sisters, where γ is either the X^0 or the largest $\overline{X}$ that α projects.

The Projection Principle now correctly states where θ-roles are assigned.

This leaves us with the job of determining which θ-roles get assigned to which positions. The θ-roles assigned to complements can be (largely) distributed by way of the c-selection. If a verb assigns two object θ-roles, like, for instance, *put*, those θ-roles can be distributed on the basis of the categories

of the phrases that bear them. We might represent this information as a list associated with the verb, as in (81).

(81) a. *put*: Agent, NP , PP
| |
Theme Location

b. *give*: Agent, NP , PP
| |
Theme Goal

The θ-role that is not associated with a c-selection, the Agent θ-roles in (81), could be understood as assigned to Specifier of $\overline{\text{V}}$. This θ-role is often called the "external θ-role." Together with the Projection Principle, this will come close to matching heads with their arguments correctly.

It's easy to see that there are some inadequacies to this system, however. It does not produce a flawless way of distributing θ-roles to complements when there is more than one complement of the same morphosyntactic category, as in (82).

(82) Jerry talked to Bill about Mary.
Jerry gave Sally the books.

Moreover, there is some evidence that the θ-role assigned externally is not completely determined by the verb alone. Marantz (1984) notes that the specific content of the subject's θ-role varies as a function not just of the verb, but also of the verb in combination with the material that follows it. This can be seen by considering (83).

(83) a. She should make a cabinet.
b. She should make an effort.
c. She should make some noise.
d. She should make nookie.

This doesn't seem to be true of the θ-role complements bear, however. It isn't possible to find examples where the specific content of the object's θ-role varies as a function of the contribution the subject makes. We may need a method of seeing the semantic combination of subject and $\overline{\text{V}}$ that takes into account more than just the meaning of the head verb. We'll return to these issues in Chapter 6. In the meanwhile, the method of thinking about θ-role assignment that (81) represents will serve.

3.5 Argument Movement

The derived subjects hypothesis conjectures that subjects of garden variety finite IPs are spoken in positions different from their D-structure position. Their D-structure position is conjectured to be the Specifier of the VP belonging to that IP, a position that the existence of small clauses indicates is a possible position for subject arguments. It relies, then, on the existence of a mechanism that gives arguments a different position than their θ-marked one, and we have introduced Argument Movement as that mechanism.

There is independent support for the existence of a mechanism like Argument Movement. There are situations in which we are more confident about where the relevant θ-marked position is, and the argument bearing that θ-role is spoken somewhere else.

3.5.1 Passive

Consider, for instance, the pair of examples in (84).

(84) a. He will make [her angry at me].
b. She will be made [angry at me].

The θ-role that *she* bears in (84b) is the same one that it bears in (84a), and yet *she* shows up in a different Case and in a different position in (84b). The VPs are different in these examples. The verb in (84b) is said to be in the "passive voice," while the verb in (84a) is in the "active voice." Passive verbs in English are embedded under the auxiliary verb *be*. It is a fairly reliable generalization about the Passive in Indo-European that Passive verbs do not assign Accusative Case, even if the Active verbs they are related to do. English conforms to this generalization. We won't look at Passive in detail, but an informal way of thinking about it is as an operation that derives a passive predicate from an active one by way of the changes described in (85).

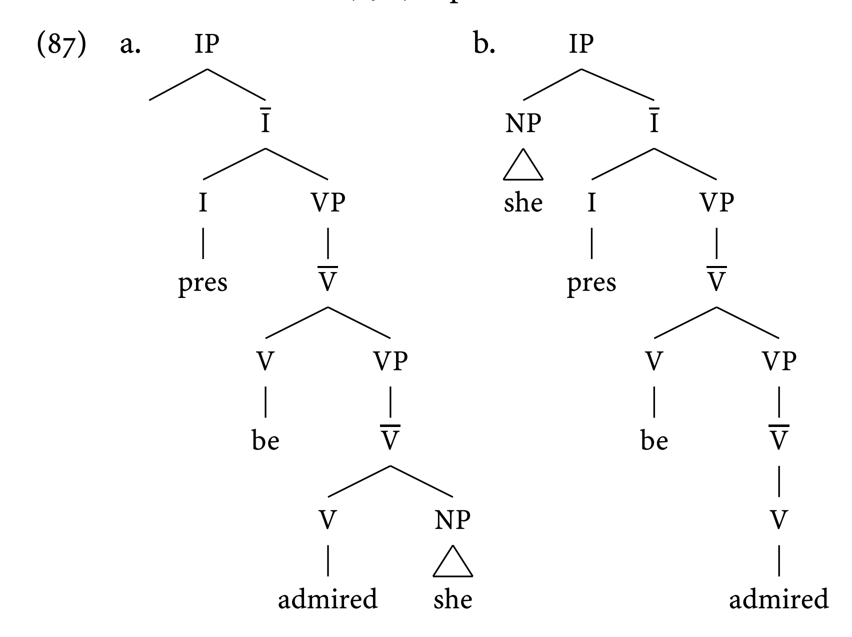
(85) PASSIVE
a. Add passive morphology to the active verb and embed it under *be*, and
b. Rob the verb of its Accusative Case, and
c. Rob the verb of its external θ-role.

This takes a garden variety "transitive" verb and produces a verb which looks "intransitive," as indicated in (86).

(86) Jerry admires her. → She is admired.

The D-structure representation of the Passive sentence in (86) is (87a) from which the S-structure in (87b) is produced.

(87) a. b.

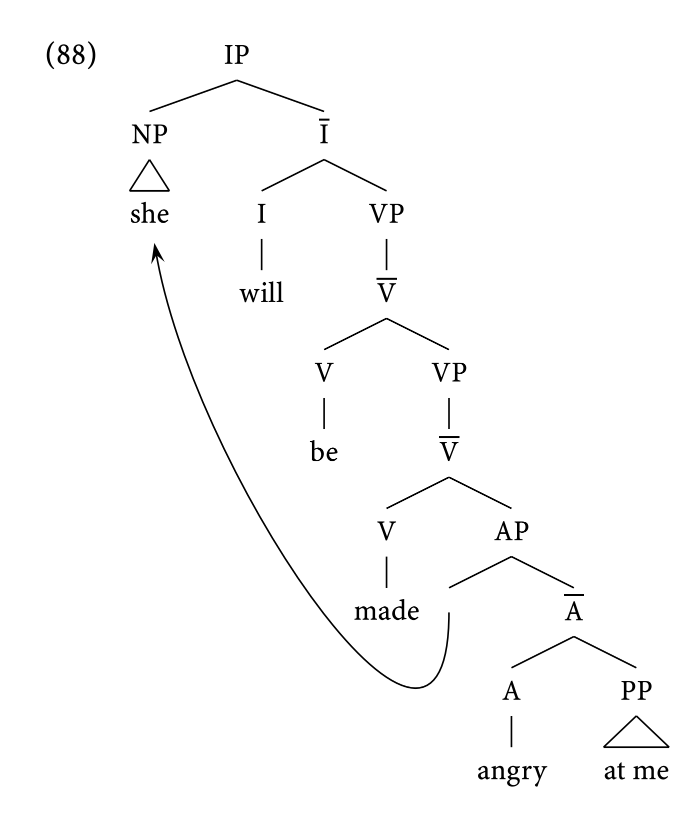


Because Passive has robbed *admired* of its ability to assign Accusative Case, its object will seek out the Nominative Case marked position to satisfy the Case filter. This will cause (87a) to be mapped onto an S-structure in which *she* is positioned in the Specifier of IP.

The passive in (84b) has the derivation in (88) on the following page. This mirrors the kind of derivation that the derived subjects hypothesis gives to every finite clause. The subject argument is put into a D-structure position from which its θ-role derives, but cannot remain there and satisfy the Case filter. Argument Movement kicks in and gives it a position in which it is Case marked. Passive Constructions, then, are ones in which we have independent support for displacement of arguments from their θ-marked positions. They even manufacture cases which resemble the situation the derived subjects hypothesis claims to be widespread.

There's a confusing aspect to English morphology which clouds how the Passive works in English. English has a morphological rule by which transitive verbs can be converted into adjectives, and these adjectives look just like Passive participles. This process is responsible for created the adjective in (89), for instance.

(88)



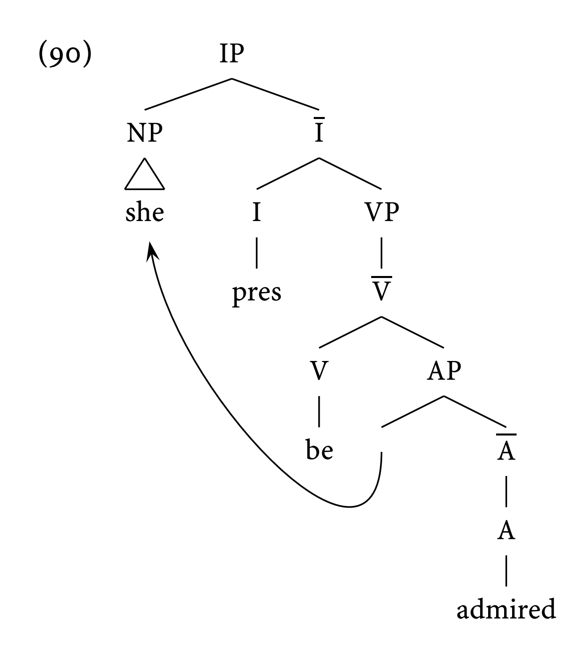
(89) the admired student

It could be, then, that what we have in (87b) is an adjectival phrase headed by *admired* rather than a Passive VP. It might have a derivation like that in (90) on the next page. Over a large class of cases, then, we will not be able to tell whether we are looking at a Passive sentence or a *be*+AP (a so-called "copular construction") sentence.

In some cases, however, it is possible to tell.[6] This is because the morphological process that derives adjectives from verbs operates on the arguments of the verb. In particular, it makes the θ-role that a verb assigns to the NP it c-selects become the external θ-role of the resulting adjective. For instance, the adjective in (90) assigns to its Specifier position the same θ-role that the verb it was derived from assigns to its complement. This is quite general. We can see this by considering the relation these adjectives have to the nouns they mod-

6 This point is made in Wasow (1977), where the way of resolving the problem we will adopt is also proposed.

(90)



ify. The pre-nominal adjective in (89) is related to the noun it modifies in the same way that the verb it derives from is related to its object. When we figure out how the modification relation works semantically, we will need to derive the fact that the modification relation for adjectives roughly matches the relation these adjectives have to their subjects in the copular construction. We can use this correlation, then, to see which θ-role is "externalized" by the adjective formation process. What we find when we look is that, in fact, it's always the "direct object" θ-role that gets externalized. In each of the cases in (91), for instance, the modified noun is related to its adjective in the same way that the direct object is to the verb.

(91) a. Sally introduced the man to Sean.
the introduced man

b. Sally placed the book on the table.
the carefully placed book

c. Sally baked the cake for Sean.
the baked cake

In no case is this relation like the one that holds between the verb and its subject argument or its "indirect object" argument.[7]

7 For more details about this process, see Levin and Rappaport (1986).

Because of this restriction on the adjective formation process, we can be certain that the example in (84b) is a Passive and not an adjective. This is because the argument that shows up as the nominative argument in this example is not an argument of the verb at all. It is the subject of the complement of the verb. If *made* were an adjective in this example, then it's the complement itself that would have appeared as subject, as can be confirmed by considering alternations like those in (92).

(92) a. Someone made this table
b. this (recently) made table.

c. Someone made him sad.
d. * the (recently) made man

3.5.2 Raising

There are verbs which behave much like the Passive verb in (84b), but without needing to be Passivized. An example is (93).

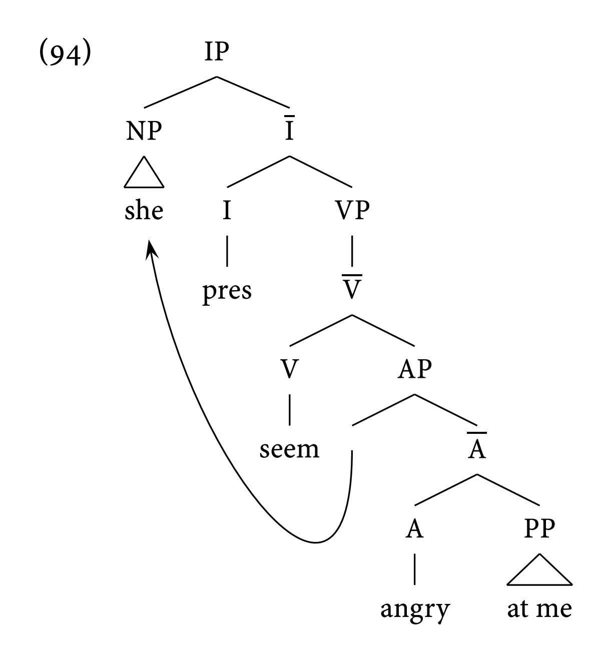
(93) She seems [angry at me].

Just as in the Passive example, *she* is the subject of *angry at me*, not *seems*. And yet, just as in the Passive, the subject of *angry at me* shows up in a higher Specifier of IP and in the nominative Case assigned to this position. This example seems to require the derivation in (94) on the facing page. Verbs like *seem* are called "raising" verbs. They have properties that force an argument in their complement to "raise" into their subject position. Under the proposals we are examining, one of those properties is that it does not assign accusative Case. Raising verbs, then, are like *be* in being among those verbs that do not have an accusative Case associated with them. As a result, in situations like (94) the argument of the embedded small clause must Argument Move into a position where it may satisfy the Case filter. In these circumstances too we see cases like those the derived subjects hypothesis claims are common.

3.6 Quantifier Float

The derived subjects hypothesis is, presently, a kind of null hypothesis. It rests on the observation that the existence of small clauses indicates that external θ-roles can be assigned to Specifier of VP, and so the null hypothesis is that they

(94)



always are. The existence of movement Transformations, or some other means of overcoming violations of Contiguity (i.e., resolving mismatches between constituent structure and linear order), makes it understandable why the surface position of the subject of finite clauses is not the Specifier of VP. And the Case filter provides a reason why a movement transformation would be invoked in these contexts: it gives an account of why subject arguments remain in Specifier of VP sometimes, and move others. The derived subjects hypothesis, then, fits together independently needed pieces — movement transformations, the Case filter, and a Projection Principle that produces small clauses — into a simple picture of how subject arguments are placed. Nonetheless, it would be encouraging to find empirical support that directly speaks on behalf of the derived subjects hypothesis.

Sportiche (1988) argues that the phenomenon of "Quantifier Float" provides such evidence. Quantifier Float arises when a quantificational term is not in the determiner position normally reserved for quantifiers, but is instead to the right of NP it is associated with. It can be illustrated by the English and French examples in (95).[8]

8 Miyagawa (1989) provides a similar analysis for Quantifier Float based on Japanese data.

(95) Les enfants (*tous) ont (tous) vu ce film (tous).
The kids (all) have (all) seen this film (*all).

Sportiche argues that there is a violation of Contiguity in these examples of the sort that movement Transformations are designed to account for. He suggests that the floated quantifier and the NP it is related to start out together at D-structure and then are separated at S-structure by Argument Movement. Note that there are language particular differences in this process that are reflected in the differing grammaticality judgments for some of the French and English examples. We'll return to these differences in the next chapter.

Sportiche's argument relies on the observation that the relation between the Floated Quantifier and the argument it is related to mimics the relation that Argument movement is designed for. Once this is appreciated, the fact that the floated quantifier appears in just those positions the derived subjects hypothesis predicts it should provides the evidence we are looking for. To see the force of his argument, we will first have to learn some generalizations about Argument Movement.

One of those generalizations is that the moved phrase can only be relocated to a c-commanding position.

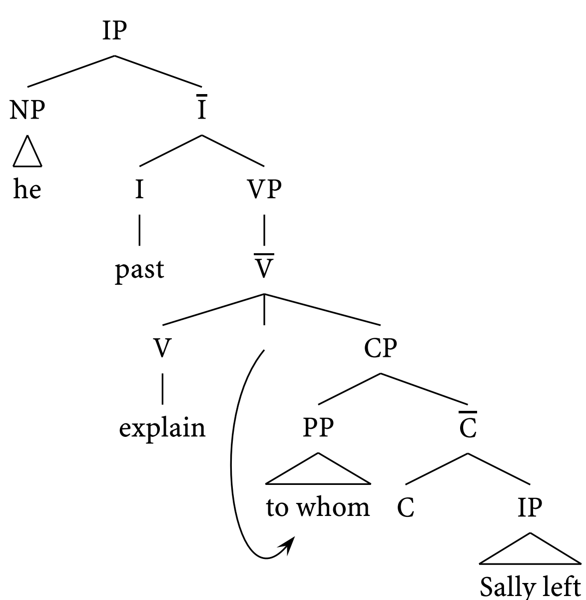
(96) If α moves to the position β, the β must c-command α.

This is a generalization about all movement Transformations: all clear cases of a term moving to a non-c-commanding position are ungrammatical. The sentence in (97) on the next page, for example, is ungrammatical because the rule of Wh Movement has moved the wh-phrase to a Specifier of CP that does not c-command the position the wh-phrase occupies at D-structure.[9] Heavy NP Shift also obeys (96), as the ungrammaticality of (98) on page 110 demonstrates. These are a small sample of examples that the movement Transformations we've defined would create ungrammatical S-structures because they do not conform to the generalization in (96). We should want to understand why (96) is true. At present, though, it does appear to be true and so we can conceive of it as a constraint on movement Transformations generally.

Sportiche points out that Quantifier Float is subject to the parallel constraint in (99).

9 I have placed the CP complement to *explain* inside the $\overline{\text{V}}$ in (97), but we have now concluded that CPs are Heavy NP Shifted to the right of VPs. I leave out this detail here just to simplify the derivation involved. This example would still violate (96) even if I'd given the correct S-structure parse.

(97) * He explained to whom Sally left.



(99) If α is a floated quantifier related to the NP β, then β must c-command α.

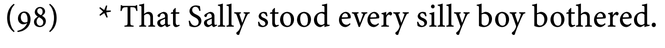
It isn't possible for (100b) to have an interpretation like that which (100a) gets, presumably because it is not possible for Quantifier Float to relate these two structures. And because (100b) can have no other interpretation – there are no other plural NPs for *tous* to be construed with – the sentence is ungrammatical.

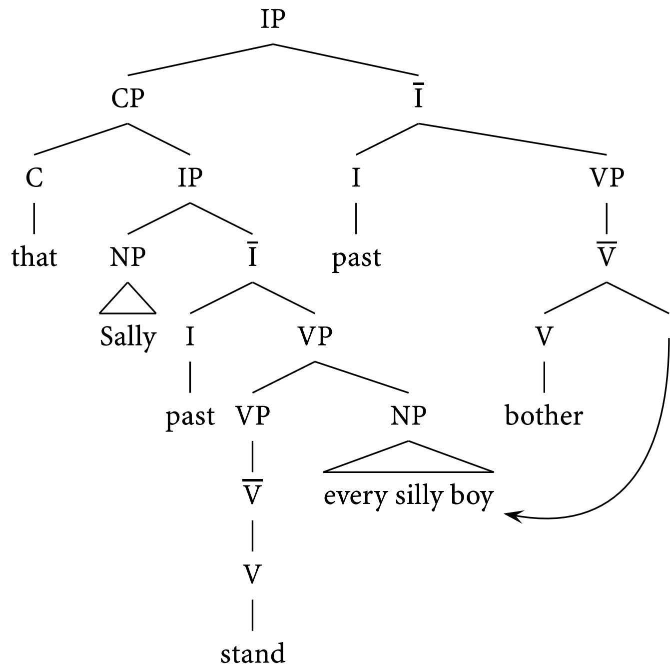
(100) a. [L'auteur de tous ces livres] a vu ce film.
the author of all these books has seen this film

b. * [L'auteur de ces livres] a tous vu ce film.
the author of these books has all seen this film

The ungrammaticality of (100b), then, can be credited to the fact *ces livres* fails to c-command *tous*.

What the c-command requirement on Quantifier Float establishes is that however the semantic connection between a floated quantifier and the NP it involves is worked out, it should follow that the NP c-commands the quantifier. There are a variety of ways that could be ensured, but because movement Transformations obey this relation, they provide one. We'll see what Sportiche's





suggestion is about the semantic connection between the floated quantifier and the NP in a moment. But first, let's consider how Quantifier Float seems to be diagnostic of Argument movement.

We have seen in the previous section that the subject argument, *the kids*, in examples like (101) has moved from a position within the small clause that follows the main verb.

(101) a. The kids seem happy.
b. The kids were made happy.
c. The kids were on the table.

And, as expected on Sportiche's proposal, a floated quantifier within the small clause can be related to *the boys*, as in (102).[10]

(102) a. The kids seem both happy.
b. ? The kids were made both happy.

10 Some speakers find (102b) marginal; I don't know why.

c. The kids were both on the table.

The relationship between the floated quantifier and the subject is like that in examples such as (103), then.

(103) a. She made the kids both happy.

b. She found the kids both on the table.

In (103), unlike (102), the NP is able to remain within the small clause and satisfy the Case filter.

Consider now (104).

(104) The parents found the kids both on the table.

This is grammatical, and has an interpretation in which *both* is construed with *the kids*. It doesn't have an interpretation, however, in which *both* is construed with *the parents*. Unlike (102), the floated quantifier in the small clause of this example cannot be construed with the argument in the higher subject position. The difference between (102) and (104) is that the argument in the higher subject position is θ-marked inside the small clause in (102), but in (104) it is θ-marked by the $\overline{\text{V}}$ containing the small clause. In fact, (104) is part of the following generalization about the relationship between a floated quantifier and the NP it is construed with.

(105) Let α be the $\overline{\text{X}}$ that assigns a θ-role to some NP. A floated quantifier construed with that NP cannot be contained in α.

Because *the parents* in (104) is θ-marked by some phrase that contains *find*, the floated quantifier is necessarily contained within this projection and (105) prevents it from being construed with *the parents*. Sportiche's proposal that Argument Movement relates a floated quantifier's position to the argument it is construed with derives (105). The phrases that Argument Movement will apply to must start out in their θ-marked positions; this is a consequence of the Theta Criterion. Therefore, because Argument movement brings arguments to c-commanding positions only, there will be no way for an argument to be placed within the $\overline{\text{X}}$ that θ-marks it.

It's clear, now, why examples like (95) provide evidence for the derived subjects hypothesis.

(95) The kids have all seen this film.

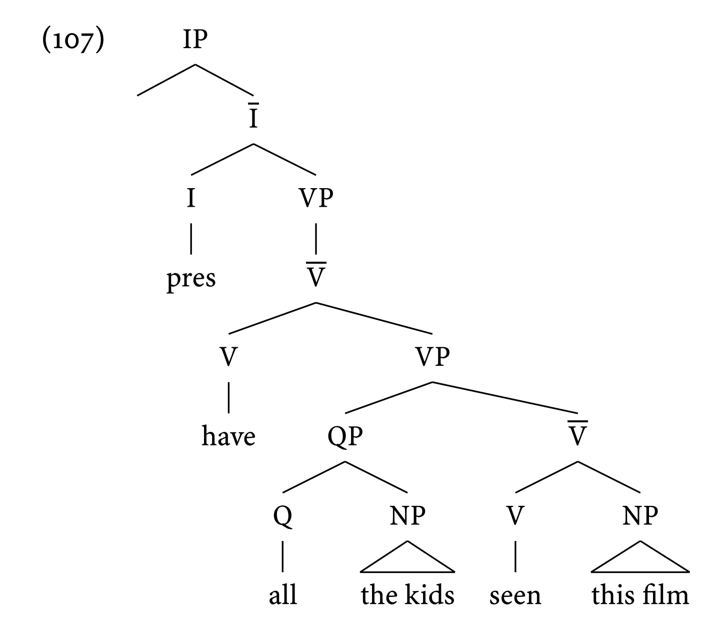
For (95) to conform to (105), the position that *the kids* gets its θ-role from must be within the phrase containing *have*. That is precisely what the derived subjects hypothesis claims.

Let's consider now how Sportiche proposes to make the semantic connection between the floated quantifier and the NP it is construed with. His leading idea is to use whatever semantic rules of interpretation derive a meaning from (106) to give what appears to be the same meaning for (95).

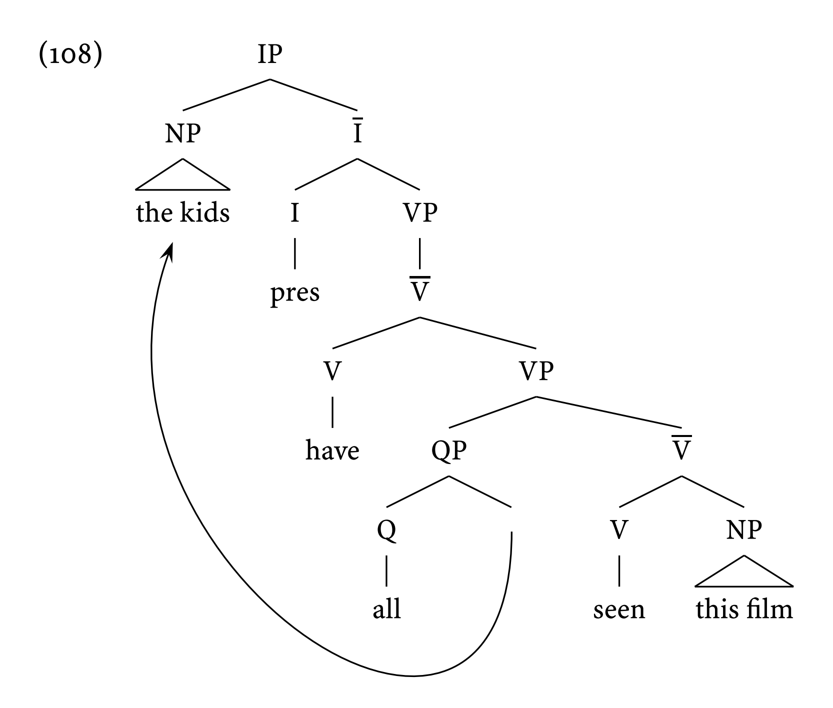
(106) All the kids have seen this film.

Imagine that the semantic connection between *all* and the NP it combines with in (106) is established at D-structure. This connection, then, is produced from the representation in (107).

(107)



This D-structure yields the S-structure representation for (106) when Argument movement relocates the entire QP *all the kids* into the Specifier of IP. It yields the S-structure for (95) when Argument movement applies to just the NP part of this QP, leaving the quantifier behind, as in (108).

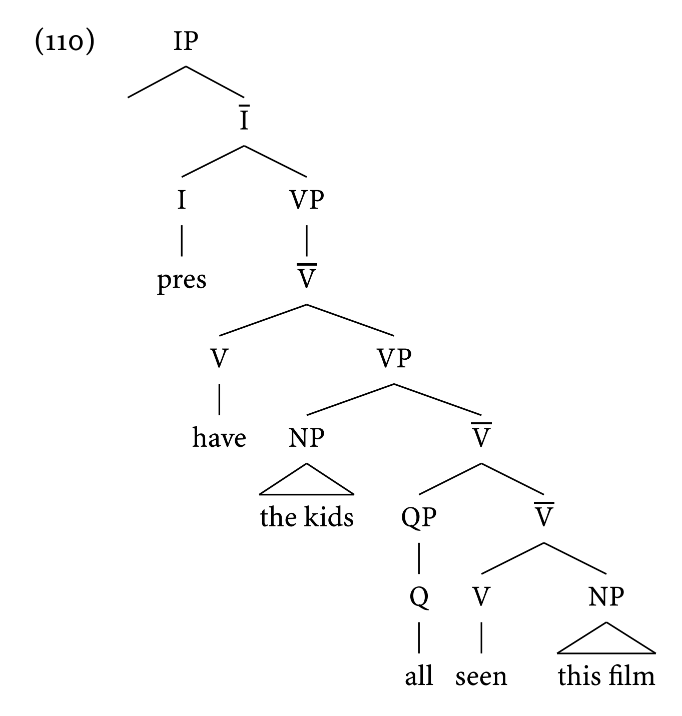
(108)


There's a popular alternative to Sportiche's account which you should keep in mind. I'll introduce it now, and offer a problem for it, but keep it at hand because it may begin to look more appealing when we discover some of the problems for Sportiche's analysis. This alternative collapses the conditions on so-called "subject oriented" (see Jackendoff (1972)) adverbs with those of Q-float. Note that these adverbs have a distribution very much like floated quantifiers.

(109) The kids (deliberately) have (deliberately) seen (*deliberately) this film (*?deliberately).

To get the interpretation right, we would need a semantics that lets a floated quantifier 'distribute' the predicate that follows it over the parts of the subject. This approach to Q-float has many advocates: Bowers (1993), for example, and Kayne (1975); (Roberts 1987, chapter 3) discusses the semantics. This alternative can be wedded to Sportiche's proposal that Argument movement is involved; it can be seen as a different way of producing the semantic composition. If we are to maintain Sportiche's suggestion that Argument movement is implicated, we should see the D-structure representation for (95) as being something like (110).

(110)



The similarities in the relationship that Argument movement produces between θ-marked position and surface position of an argument and that Quantifier Float seems to require between a floated quantifier and the NP it is construed with recommend maintaining this part of Sportiche's proposal. What we might consider altering is how the quantifier combines syntactically with the NP it is construed with. Sportiche's (best) argument against this alternative is that the distribution of subject-oriented adverbs and floated Q are not always the same cross-linguistically.

There's another kind of argument against this alternative which is found in Giusti (1990b,a) which is worth looking at. Her argument is typological, and based on where floated quantifiers can be found related to objects across languages. First, note that floated quantifiers related to objects are, by themselves, something of a puzzle for an approach like the alternative sketched above. That account demands that the floated quantifier combine with a $\overline{\text{V}}$, or similar predicate, and steer how that predicate combines semantically with the neighboring NP. The existence of floated quantifiers related to objects, then, is by itself an argument for Sportiche's imagine of the underlying connection between quantifier and NP. What Giusti observes is that the languages which allow quantifiers to be related to objects are those for which we have independent evidence that objects can move.

Thus, German/Dutch and Icelandic, but not English or Danish, can have a floated quantifier related to an object, as in (111) and (112).

(111) a. Der Lehrer hat die Schüler alle gelobt.
the teacher has the students all praised
'The teacher has praised all the students.'
(German)

b. De leraar heeft de kinderen allen geloofd
The teacher has the children all praised.
'The teacher has praised all the students.'
(Dutch)

c. Stúdentarnir lasu greinina allir.
students-the read article-the all
'The students read all the articles.
(Icelandic)
(Giusti 1990a, (1): 137, (22a) & (25b): 144)

(112) a. * The teacher has praised the children all.

b. * Laereren roste eleverne alle
teacher-the praised children-the all
'The teacher praised all the students.'
(Danish)
(Giusti 1990a, (2c): 137)

She argues that this can be related to the fact that definite objects in Dutch, German and Icelandic can move leftwards, whereas this is not possible for objects in English and Danish. That leftwards movement is possible in these languages is, at least partly, indicated by the fact that they can precede adverbs which are normally found at the left edge of VPs in these languages. We have the contrasts, then, between examples such as (113) and those in (114).

(113) a. Der Lehrer hat die Schüler ohne Zweifel gelobt.
the teacher has the students without a doubt praised.
'The teacher has without a doubt praised the students.'
(German)

b. De leraar heeft de kinderen gisteren geloofd.
the teacher has the children yesterday praised.
'The teacher has yesterday praised the students.'
(Dutch)

c. Stúdentarnir lasu greinina ekki.
students-the read articles-the not
'The students didn't read the articles.'
(Icelandic)
(Giusti 1990a, (2): 136, (22a): 144, (13): 141)

(114) a. * The teacher has praised the children not.
b. * Laereren roste eleverne uden tvivl.
teacher-the praised students-the without a doubt.
'The teacher without a doubt praised the students.'
(Danish)
(Giusti 1990a, (2): 137)

This is just what we would expect on a Argument Movement account of Quantifier float, of course, because on that account movement is what is responsible for separating an NP from the quantifier it is construed with. But this is also an argument for Sportiche's particular way of putting the quantifier together with the NP. On Sportiche's proposal, the only way a quantifier can appear after an NP it is construed with is by that NP's movement. The alternative account, by contrast, allows the D-structure arrangement of quantifier and NP to put the quantifier in a position that follows the NP. It's not clear how such an account could capture Giusti's correlation.

On the other hand, there are examples which Sportiche's proposal does not explain. As he notes, his model of Quantifier Float should allow quantifiers to be related to objects in English in just those cases where the object has undergone Argument movement. We've seen that one situation where this happens is in the Passive: an object moves into a higher Case-marked position when the verb it is an object of is Passivized. And yet, Quantifier Float is not possible in these contexts, as the ungrammaticality of (115) indicates.

(115) * The kids were praised all.

For some steps towards a solution, see Bošković (2001, 2004).

Another problem concerns examples like (103).

(103) She made the kids both happy.

In this example, *the kids* does not undergo movement. Its D-structure position is also the one where it gets assigned Case. If quantifiers follow the NPs they are construed with only when they move, how is it that the quantifier manages to follow the NP in this example? We will see a solution to this problem in Chapter

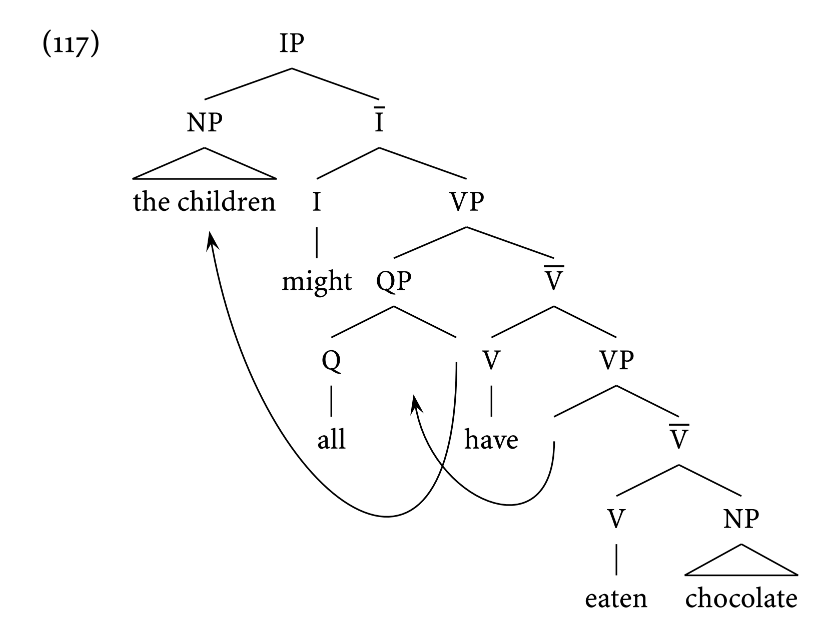
6. Keeping these problems in mind, let's nonetheless adopt Sportiche's account of Quantifier Float.[11]

One interesting consequence of Sportiche's analysis is that floated quantifiers can be seen as indicators of the positions that arguments occupy not only in the D-structure element of the derivation, but in all the non-S-structure parses of the derivation. It gives us, in other words, a useful glimpse at these otherwise invisible parses. We learn from this that phrases can take smaller steps than necessary in their journey from a θ-marked position to a Case marked position. So in (116), for example, it appears that *the children* has made an intermediary stop in the Specifier of the VP headed by *have*.

(116) The children might all have eaten chocolate.

Argument Movement is employed twice, as indicated in (117).

(117)



We learn from these sorts of examples that it's possible for Argument movement to act on a structure that is itself the product of Argument movement. Such derivations are said to involve "successive cyclic" movement. The existence of such derivations means that complex interactions of movements are conceivable, and, as a result, the mapping from D-structure to S-structure is

11 For other work on Quantifier Float, see Jaeggli (1982), Belletti (1982), Déprez (1989), Doetjes (1995) and Bobaljik (2003).

potentially quite complex. There are problems that lurk here, but they are for another time.

3.7 Infinitives

We have seen that "raising verbs" select small clauses but do not assign Case to the NP argument within that small clause. As a result, Argument movement brings that NP argument into a Case marked position. In (118), for instance, *Sally* has moved from within the small clause where its θ-role is assigned, into its surface, nominative Case marked, position.

(118) Sally seems happy.

Raising verbs are (some of them) also capable of selecting an IP headed by *to*, as in (119).

(119) Sally seems to be happy

In this example too, *Sally* has moved from a position within the small clause into its surface, nominative Case marked, position.

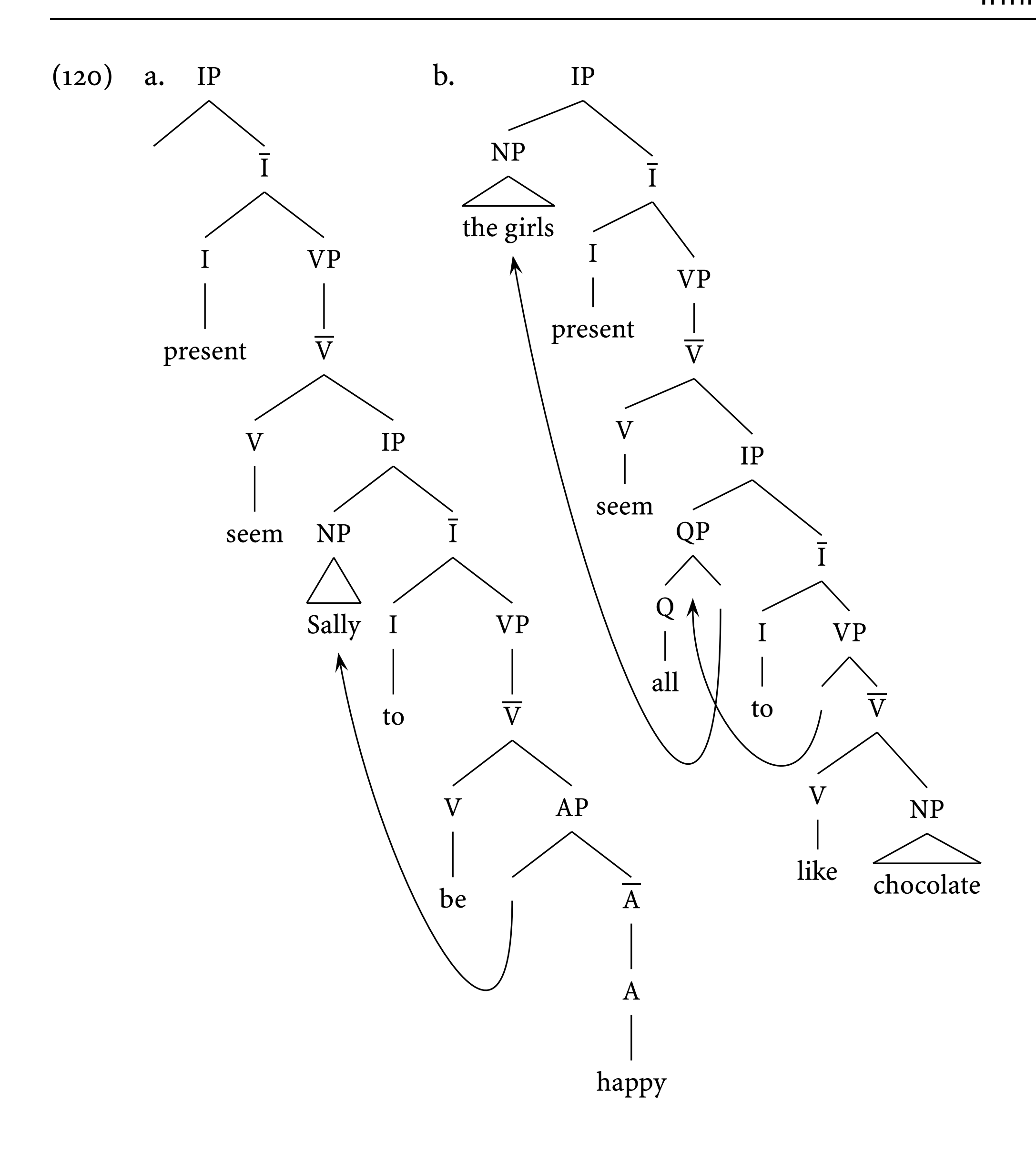
But in this case, there is a question about how the EPP is satisfied. (The EPP, recall, is the requirement that a Specifier of IP be filled with something.) It's clear how the highest IP satisfies the EPP: its Specifier is filled with the subject argument that resides in it at S-structure. But what of the Specifier of the IP headed by *to*? One possibility is that (119) has a successive cyclic derivation in which a representation like that in (120a) on the next page is manufactured. From this representation, A-movement could apply again to form the S-structure we see in (119). We might imagine, then, that (119) has a successive cyclic derivation. If this is correct, we should understand the EPP to be a condition that can be satisfied by any phrase marker in a derivation.

That a successive cyclic derivation of this sort is possible is indicated by (121), whose derivation is indicated in (120b) on the facing page.

(121) The girls seem all to like chocolate.

We see in this example that Argument movement has moved *all the girls* into the Specifier of the infinitive before the NP part is moved into its surface position. The presence of the floated quantifier in the Specifier of the infinitival IP is the tell-tale sign that the subject argument has passed through the Specifier of the infinitival clause.

(120) a. b.



Infinitival clauses come in a variety of types that can be distinguished from those that following raising verbs, and we will look at some of these in this section. The example in (122), for example, though it has a superficial similarity to (119), is different in some ways.

(122) She tried [to be happy]

In (119), the count of θ-roles and arguments matched, as expected. But in (122), this isn't the case. Both *tried* and *angry* assign an external θ-role, and yet there is only one argument which seems to bear them both: *she*. This is precisely what the Theta Criterion prohibits. Recall, that the Theta Criterion is designed in this way in an attempt to capture the fact that (123a) doesn't have the meaning that (123b) does.

(123) a. I showed John.
b. I showed John himself.

So, either we should find another explanation for the fact that (123) illustrates, or we should find a way of resolving (122) with the Theta Criterion.

Note also that (122) seems to be a violation of the Extension to the Projection Principle. Unless we determine that (122) involves Argument movement of *she* through the Specifier of the infinitival IP, as in the raising case, there appears to be nothing that occupies this Specifier position at any point in the derivation. Moreover, it's not just that (122) seems to counterexemplify these principles, it perversely can't satisfy them. If the Specifier of the embedded IP is given content, and in so doing hosts an argument for the θ-role that *angry* assigns, the result is ungrammatical.

(124) * She tried [him to be angry].

And if a predicate that does not have an external θ-role to assign is put into these infinitives, the result is also ungrammatical.

(125) * She tried [to seem that pigs fly].

Infinitives with this peculiar mix of properties are called control infinitives, and a wide assortment of verbs select infinitives of this type. Some are in (126).

(126) a. Sally prefers to eat chocolate.
b. Sally managed to eat chocolate.
c. Sally attempted to eat chocolate.
d. Sally wished to eat chocolate.
e. Sally asked to eat chocolate.
f. Sally proposed to eat chocolate.
g. Sally neglected to eat chocolate.
h. Sally promised to eat chocolate.
i. Sally decided to eat chocolate.

In all of these examples, there is an argument — here it is *Sally*, the subject of the higher verb — that appears to bear the θ-role assigned to the subject of the infinitival clause. We say in these cases that the subject argument "controls" the infinitival clause, or "controls" the embedded subject. There are instances of control infinitives where it's the object of the higher verb that is the controller; (127) provides some examples.

(127) a. Sally told Jim to eat chocolate.
b. Sally asked Jim to eat chocolate.
c. Sally required Jim to eat chocolate.
d. Sally forced Jim to eat chocolate.

In these examples, it is the object, *Jim*, not the subject, *Sally*, that is understood to bear the external θ-role assigned by *eat*.

From what we have seen about control infinitives, we can draw the following conclusions.

(128) a. A control infinitive must have a subject argument.
b. The subject argument of a control infinitive cannot be pronounced within the infinitive.
c. The subject argument of a control infinitive can be understood as the same as an argument in the higher clause: its "controller."

Our goal is to understand what is responsible for these properties of control infinitives. Our hope should be that it is consistent with, maybe even follows from, the forces that are responsible for positioning arguments. We'll look at a few ideas in this section that might bring us within range of these desires. I'll start with the question of whether these infinitives are CPs or not, and then turn to issues that more transparently bear on (128).

3.7.1 The CP/IP status of infinitives

Despite the lack of a complementizer, there is some evidence that control infinitives can be a CP. Of the verbs that select a control infinitive, those that also s-select a question can combine with a control infinitive that begins with a wh-phrase; (129) on the next page are some examples. Recall that wh-phrases are moved into the Specifier of CP by Wh movement. We should conclude, then, that the infinitives in (129) are CPs. And in these very contexts, the complementizer *whether* is, as expected, also possible as (130) shows.

(129) a. Sally told Jim how to eat chocolate.
b. Sally asked when to leave.
c. Sally decided when to leave.
d. Sally wondered what to eat.

(130) a. Sally told Jim whether to eat chocolate.
b. Sally asked whether to leave.
c. Sally decided whether to leave.
d. Sally wondered whether eat chocolate.

In many, not too distantly related, languages, even the non-question form of a control infinitive comes with a complementizer. Icelandic is a particularly clear example of this. It has subject and object control infinitives that closely resemble those of English. And yet these control infinitives appear with the same complementizer that is found in finite clauses in Icelandic: *að*.

(131) María lofaði **að** lesa bókina.
Mary promised that read the book
'Mary promised to read the book.'

compare:

María segir **að** þú hafir lesið bókina.
Mary says that you have read the book
'Mary says that you have read the book.'

We should conclude, then, that a control infinitive can be a CP.

By contrast, the same kind of evidence regarding raising infinitives suggests that they are never CPs. There are no known examples of raising infinitives that have the syntax of questions.

(132) a. * It seems whether to have left.
compare: John seems to have left.
b. * It appears when to eat chocolate.
compare: John appears to eat chocolate.
c. * It is likely who to dance.
compare: John is likely to dance.

This could be because infinitival clauses are incapable of being CPs, and therefore do not support the CP-dependent syntax of questions. But it could also be because raising verbs (and adjectives, such as *likely*) do not semantically select

questions. Whether a verb can take a question as an object depends on the semantics of the individual verb. A verb like *believe*, for instance, cannot take a question complement, whereas the similar *know* can.

(133) a. * Jill believes whether Sean has left.
b. * Jill knows whether Sean has left.

The absence of question complements to raising verbs could, then, simply reflect a fact about the semantics of these predicates, and nothing about the syntax of their infinitival complements. Indeed, raising predicates do not seem able to combine with question complements even when those complements are finite clauses.

(134) a. * It seems whether Joe has left.
compare: It seems that Joe has left.
b. * It appears when María has eaten the chocolate.
compare: It appears that María has eaten the chocolate.
c. * It is likely who will dance.
compare: It is likely that Anna will dance.

Finite clauses can always be CPs (it seems), and so we should conclude that the absence of question complements to raising verbs has to do with their meaning.

Although we do cannot conclude that raising infinitives are prevented from being CPs, we also cannot learn from their inability to be questions whether they are allowed to be CPs. Icelandic, however, provides a clue. Raising infinitives in Icelandic cannot include the complementizer *að*.

(135) * María hafði virst að hafa vaskað upp diskana.
Mary had seemed that have washed up the dishes
'Maria had seemed to have washed up the dishes'

compare:

María hafði virst hafa vaskað upp diskana.
Maria had seemed have washed up the dishes
'Mary had seemed to wash up the dishes.'

This would be explained if raising infinitives are IPs, but cannot be CPs.

In Icelandic, then, we can conclude that one difference between raising and control infinitives is their categorial status. Raising infinitives are IPs but cannot be CPs, whereas control infinitives are CPs. In English, the picture isn't as

clear. There is no direct evidence that raising infinitives cannot be CPs. And though there is evidence that some control infinitives are CPs — the questions — other control infinitives do not differ outwardly from raising infinitives. English internal evidence supports the conclusion that control infinitives can be CPs, but nothing more.

If we consider the evidence from Icelandic and English together, however, and keep in mind the criterion of explanatory adequacy, we can find some help in making decisions about the proper treatment of English infinitives. We conclude from Icelandic that the mapping in (136) exists in at least one grammar.

(136) raising infinitive → IP
control infinitive → CP

Assuming that this is not a capricious relationship, we can assume that there are forces, still to be discovered, that produce it. If the mapping in (136) does not materialize in English, then we can conclude that those forces do not operate in the same way in English. When we model these forces, our model should allow them to vary from language to language. Consequently, we must credit the child with the ability to detect that variation as that child is acquiring his or her language. Let's call this "hypothesis 1."

Hypothesis 2 is that the forces responsible for the existence of (136) in Icelandic are also found in English. On this hypothesis, Icelandic and English should be identical with respect to the categorial status of their infinitives. Any differences in the appearance of infinitives in these languages will flow from other ways these languages differ. What are the differences between Icelandic and English that would have to be explained on hypothesis 2? As we've seen, Icelandic and English raising infinitives look the same; nothing would have to be said under hypothesis 2 about how raising infinitives differ in these languages. Control infinitives in Icelandic and English also look the same in all ways except one: declarative control infinitives can appear with a complementizer in Icelandic but not in English. Hypothesis 2 requires that this difference be made compatible with the mapping in (136). One way of doing that is with (137).

(137) The complementizers of infinitival clauses in English have no phonetic realization.

It seems difficult to avoid a conclusion like (137) under hypothesis 2. Hypothesis 2 is likely to require, then, that words can be silent. It must credit the child with the ability to detect the existence of silent words in a sentence.

So we have before us two alternatives. English does not conform to the mapping in (136) and the forces that are responsible for (136) can vary in ways that are detectable by children acquiring English. Or, English does conform to the mapping in (136) but has a silent infinitival complementizer, and the existence of this silent complementizer is detectable by children learning English. Framing the analysis of English infinitives in this way allows us to make decisions based on criteria of language acquisition. If the forces behind (136) turn out to be too distant from the information available to a child, then we can conclude that they cannot vary in detectable ways and we're stuck with hypothesis 2. If, by contrast, the existence of words in a sentence cannot be detected without their phonetic manifestation, then we're stuck with hypothesis 1. Hypothesis 2 is adopted in *Lectures on Government and Binding*, and it is, perhaps, the most popular in the subsequent literature. One way of ensuring that the silent complementizer is detectable to the language-acquiring child is to elevate the mapping in (136) to a language universal. If the child can rely on (136), then it can deduce that there is a silent complementizer in the declarative control infinitives of English. Let's adopt hypothesis 2 as our working hypothesis as well. We'll examine some alternatives in a later chapter.

Before leaving the question of whether infinitives are CPs or not, we should look at one last "kind" of infinitive. Some predicates that take control infinitives as objects can take an infinitive which does have an overt complementizer, and in which the subject argument is visible inside the infinitive. Some examples of these predicates and the two types of infinitives they can combine with, are given in (138).

(138)
a. Misato wanted to eat durian.
   Misato wanted for him to eat durian.
b. Junko preferred to eat natto.
   Junko preferred for me to eat natto.
c. It is possible to eat poi.
   It is possible for some to eat poi.

The subject of the infinitival clauses appears in the accusative Case, as is clear from the form of the pronouns in (138a) and (138b). The complementizer is *for*, which might easily be mistaken for the homophonous preposition found in sentences like *She wanted something for Susan*. We think that the *for* in (138) is a Complementizer for several reasons. First, it forms a constituent with the entire clause that follows it. This is indicated, among other things, by the behavior of these phrases in the context of cleft constructions.

(139) a. It's **for him to eat chocolate** that Sally would prefer.

b. * It's to him how to eat chocolate that Sally should explain.

The ungrammaticality of (139b) derives from the fact that there is only room for one phrase between the *it's* and the *that* of these clefts, and in (139b) two things, a PP and an infinitive, have been shoved into this spot. Thus the grammaticality of (139a) would seem to argue that we do not want to parse the boldfaced string as a PP followed by an infinitive, but instead as something that makes a single phrase. To the extent that an infinitive is an IP, and that the *him* in (138) is its subject, then one of the few ways of doing this is to let *for* be a complementizer.

Moreover, the *for* that comes with these infinitives has a property that no preposition in English has, but which the complementizer *that* does have. It is optional. Under a certain set of circumstances, the complementizer *that* can go unspoken in English, as in (140).

(140) a. Sally said that he eats chocolate.

b. Sally said he eats chocolate.

This is also possible for the *for* of those infinitives that are objects to verbs, as (141) shows.

(141) a. Misato wanted him to eat durian.

b. Junko preferred me to eat natto.

Furthermore, the circumstances on the optionality of the complementizer *that* are mimicked by the optionality of *for*. In general, the complementizer *that* can go unspoken only in contexts where the CP it heads is a complement to a verb. That is why *that* cannot go missing when it heads a CP that has been clefted, as in (142). A parallel constraint on *for* ellipsis is indicated by (143).

(142) a. It's that he eats chocolate that Sally said.

b. * It's he eats chocolate that Sally said.

(143) a. It's for him to eat chocolate that Sally would prefer.

b. * It's him to eat chocolate that Sally would prefer.

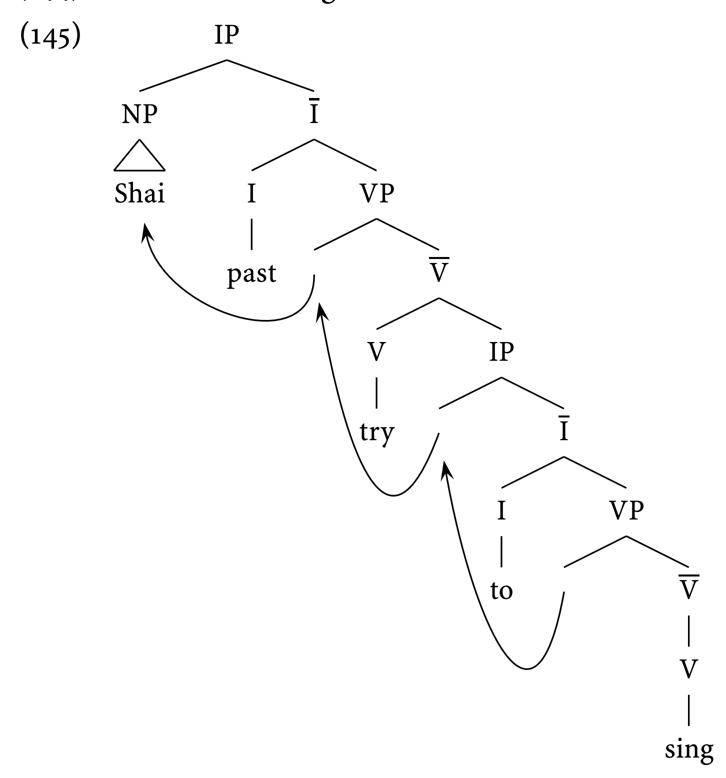
In general, the constraints on where *that* can go missing are parallel to those on where *for* can, suggesting that they are both subject to some process, restricted to complementizers, that allows them to go silent.

3.7.2 On the Control relation

Let's return now to how control infinitives express their arguments. In the cases we've looked at so far, they are rather like raising infinitives in that an argument of their main predicate surfaces in a position in the higher clause. One possibility, then, is that control infinitives use the same mechanism that raising infinitives do to put their arguments in this position. We might imagine that Argument movement relocates the argument NP of (144) in the way indicated in (145).

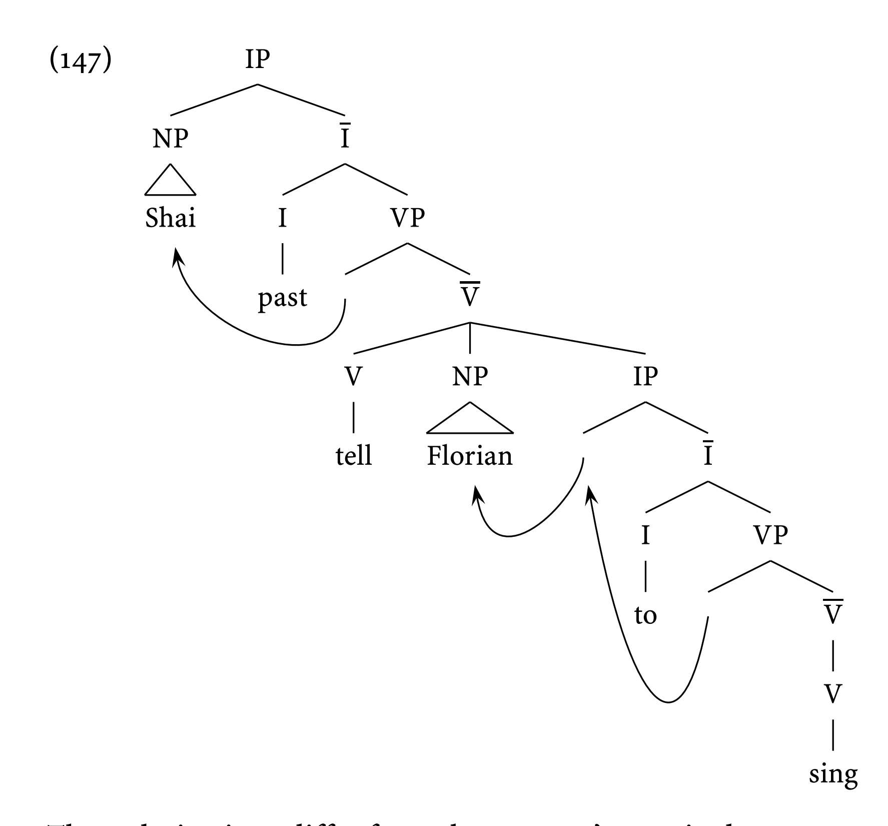
(144) Shai tried to sing.

(145)



An instance of object control, like that in (146) below could have the derivation indicated in (147).

(146) Shai told Florian to sing.

(147)


These derivations differ from the ones we've posited up to now for Argument movement in that they move arguments into θ-marked positions. If this is the correct account of control infinitives, then, we should adjust the conditions we've placed on Argument movement to allow these derivations. Doing so would not only give a transparent account of the Control relation, it would also provide an account for how control infinitives satisfy the EPP and that part of the Theta Criterion that requires every θ-role to be assigned to an argument. On this view, the Control relation is just the one established by Argument movement between an argument and its θ-marked position. This view of the Control relation is offered in Hornstein (1999) and its consequences explored in Hornstein (2000) and Boeckx and Hornstein (2003, 2004, 2006a,b).

There are circumstances where control infinitives do not involve control, however; in these circumstances the subject argument is not an argument expressed elsewhere. This is the case when, for instance, the verbs taking a control infinitive are passivized, as in (148).

(148) a. To eat natto shouldn't be tried.
b. To be content is often wanted.

c. To be honest should be promised in a wedding vow.

In fact, whenever a control infinitive stands in subject position, its understood subject does not have to be an argument found elsewhere; (149) provides some other examples.

(149) a. To remind him of Amherst Winters bothers Pius.
b. To eat durian shouldn't be possible.
c. To leave your shoes untied is a danger to those around you.
d. To be the largest prime number is to be very large indeed.

In these examples the subject of the infinitive is understood to have a meaning close to that which *one* gets in sentences like (150), as can be appreciated by comparing this example to (149b).[12]

(150) For one to eat natto shouldn't be possible.

The understood subject of these infinitives is said to be "arbitrary."

In (148) and (149) there is no argument that could be the understood subject of the infinitive and so it is easy to see that it must get another meaning. But there are also examples where potential controlling argument is present and yet an arbitrary reading for the understood subject of the infinitive is available. That is the case for the examples in (151), for example, which may all have the same arbitrary interpretation that is found in (149).

(151) a. Elena discovered that to remind Pius of Amherst Winters bothers him.
b. Jeff thought that to eat durian shouldn't be possible.
c. Bart remembered that to leave his shoes untied is a danger to those around him.
d. Jill knew that to be the largest prime number is to be very large indeed.

When the infinitive is a question, it may also sometimes support an arbitrary interpretation for its understood subject, as in (152).

(152) a. Jon knew how to eat live shrimp.
b. Satoshi can tell when to add vinegar to rice.

12 There is a subtle difference in the meanings of (150) and (149b), but it doesn't seem to me that this difference resides in the meaning of the subjects.

c. Julie decided where to put the *Dicentra*.

These can get interpretations roughly parallel to those in (153), in which the subject is arbitrary.

(153) a. Jon knew how one can eat live shrimp.

b. Satoshi can tell when one should add vinegar to rice.

c. Julie decided where one should put the *Dicentra*.

The precise conditions under which the arbitrary interpretation are available are imperfectly understood. See Williams (1980, 1989) and Landau (2004) for some ideas.

Control infinitives with an arbitrary reading pose the same problem for the EPP and Theta Criterion that normal control infinitives do. They seem to violate the EPP — there is nothing in the Specifier of their IP — and because the θ-role normally assigned to the subject has no argument to bear it, they seem to violate the Theta Criterion as well. Moreover, just like normal control infinitives, they are ungrammatical if there is no θ-role for a subject argument, as (154) shows.

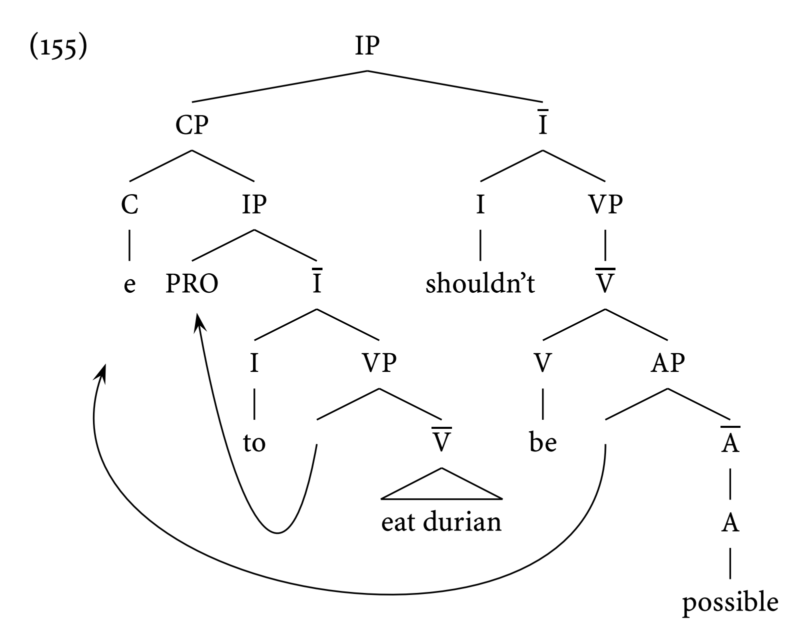
(154) * To seem that pigs fly is worrisome.
compare: For it to seem that pigs fly is worrisome.

Just as for normal control infinitives, then, these facts taken together suggest that there is a subject argument present in these infinitives. If such an argument is present, then the EPP and Theta Criterion could be maintained. Moreover, the ungrammaticality of examples like (154) would be explained: if there is a subject argument, then the Theta Criterion requires that there also be a θ-role assigned to it. In these cases, however, this alleged subject argument cannot be found in some unexpected position. We cannot here finger the controller as that argument because there is no controller.

A common reaction, and the one I shall adopt, is that there is a subject argument syntactically present in these examples but that it is silent.[13] This argument is indicated by putting "PRO" in the position our syntax suggests it should be. The representation for (149b), then, would be (155).

13 For an approach that does not countenance a silent subject, see Chierchia (1984).

(155)



I've decided that this infinitival clause is a CP with a silent complementizer, represented here with "e." I've also shown PRO moving from the position that the external θ-role is assigned to, Specifier of VP, to the Specifier of IP where it satisfies the EPP. I also show the infinitival CP itself moving from Specifier of AP into Specifier of IP; this claims that the infinitival gets an external θ-role from *possible* and moves in order to satisfy the EPP.

We can think of PRO as being a kind of pronoun (hence its name), maybe along the lines of a silent *one*. Of course, if the grammar of English includes a silent pronoun, then where it can be must be heavily constrained. Otherwise, sentences such as (156) should be grammatical, with Theta Criterion satisfied by the presence of PRO in subject position or object of preposition.

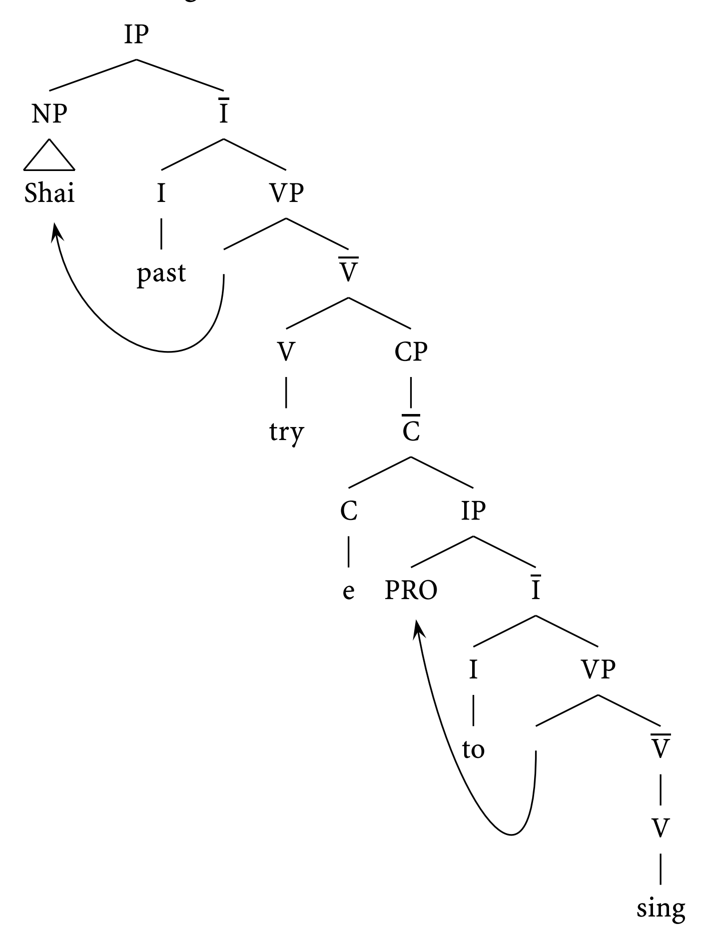
(156) a. * Ran into the garden.
b. * Andre talked to.

The problem of restricting PRO's distribution remains unsolved. In fact, the question of where PRO can be is still controversial. There is no controversy over the question of whether PRO can be found in complement positions; it seems clear that it cannot. And it also seems clear that PRO cannot be in subject position of finite clauses, at least not in English. But it's not clear yet in which non-finite clauses PRO can be a subject.

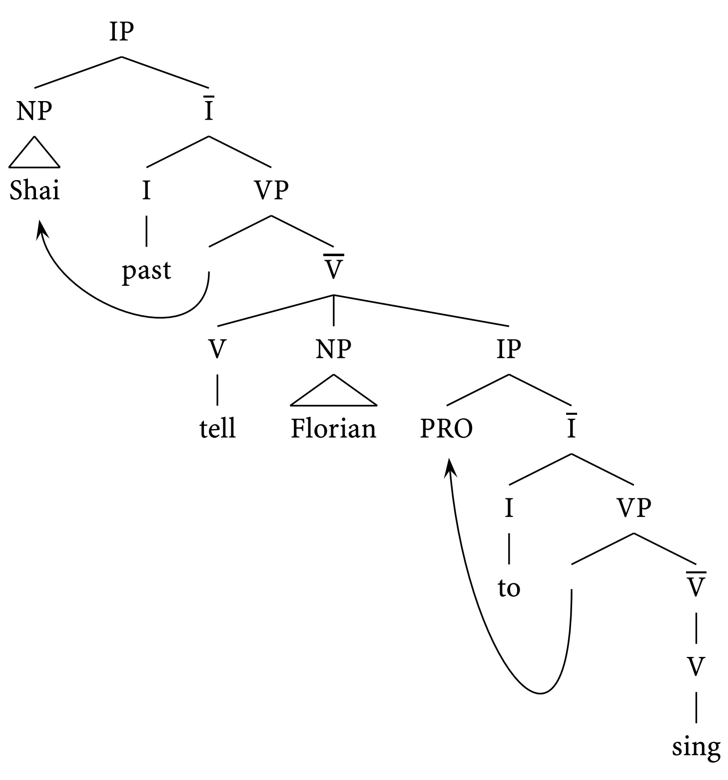
Perhaps the most popular view presently is that PRO can stand as subject to any control infinitive, not just those that get an arbitrary interpretation. On this

view, the Control it does not reduce to a special instance of Argument Movement, but is instead an interpretation that PRO can receive. The representation for sentences involving Control would not be as indicated in (145) and (147), but would instead include PRO, as in (157) and (158).

(157) Shai tried to sing.



(158) Shai told Florian to sing



Unlike the movement-based analysis of Control that (145) and (147) illustrate, the representations in (157) and (158) obey the Theta Criterion as it is now formulated. What we would need to do if these representations are correct is produce a non-movement account of Control, an account that would emerge as part of a more general account of how PRO gets its semantic interpretation.

What we've discovered in this section, then, is that there is a silent pronominal argument in English and that this silent argument is prevented from appearing in a variety of positions otherwise normally available to arguments. Keeping in mind that this argument, PRO, cannot be in complement positions, we can partially describe its special distribution with (159).

(159) The PRO Restriction
PRO can only be in certain Specifier positions at S-structure.

Which Specifier positions PRO can be in is controversial, but it certainly doesn't include the Specifiers of finite IP.

The PRO Restriction describes something that is peculiar to the placement of the silent argument: PRO. Let's now look at whether the conditions we've

been developing to capture the placement of the non-silent arguments also hold of PRO.

3.7.3 Argument movement and PRO

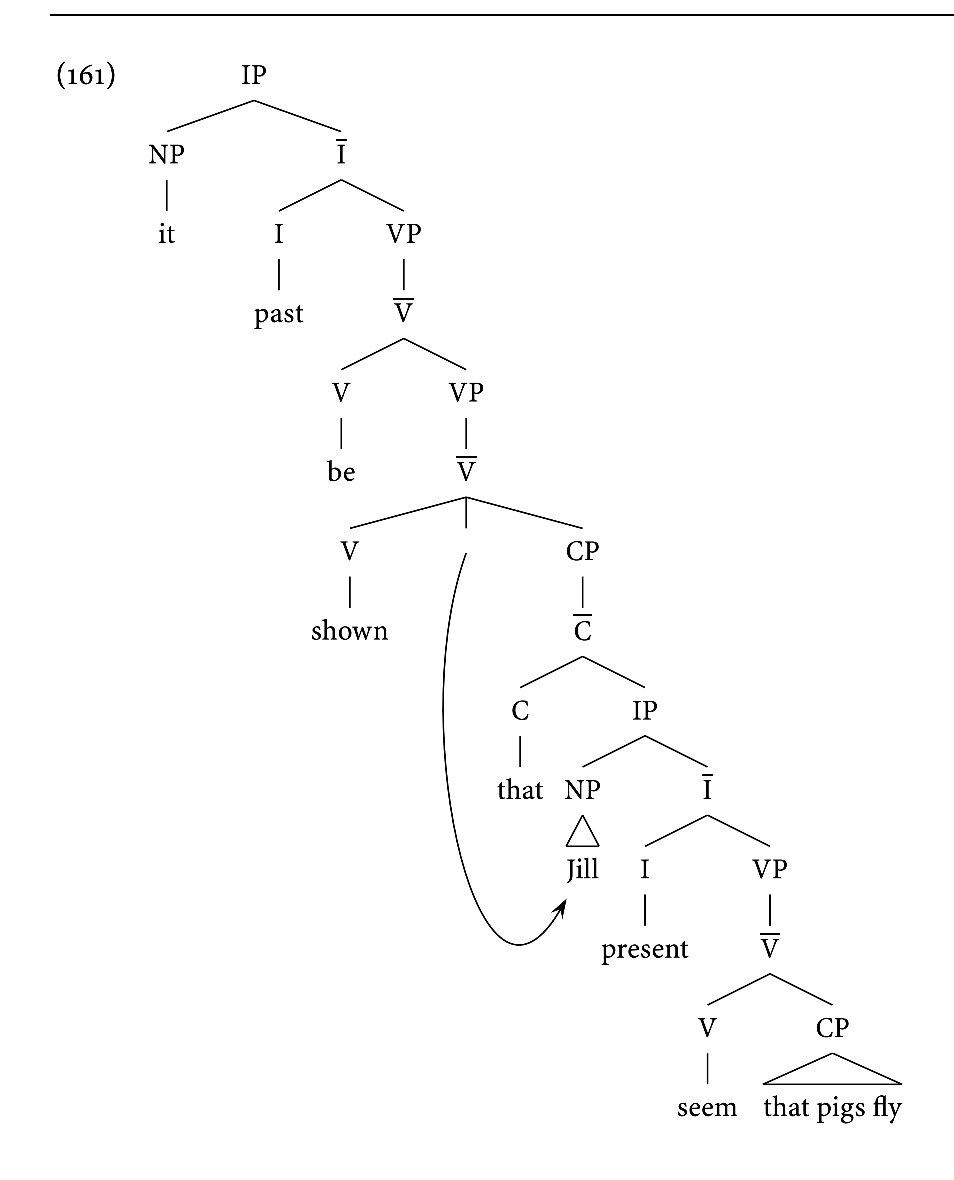
To this point, Argument movement has been invoked in situations where an NP is θ-marked in a position that is not Case marked. Because argument NPs must be both Case marked and receive a θ-role, there will be no single parse that can satisfy both of these conditions in such a case. Argument movement therefore produces a series of parses: one of which will satisfy the Theta Criterion and Projection Principle and another one of which will satisfy the Case filter. I have expressed the Theta Criterion and Projection Principle as well-formedness conditions on D-structures, and the Case filter as a well-formedness condition on S-structure. This places certain conditions on how Argument movement will arise because it imposes requirements on what the input and output parses can be. If Control is an instance of Argument movement, then certain changes to the Theta Criterion will have to be made, and there will be consequent changes to how Argument movement behaves.

In addition to the constraints imposed on Argument movement by these conditions, we have seen that there is a condition imposed directly on Argument movement, and other movement operations: the moved term can only be relocated to a position that c-commands its starting position. This condition is responsible for blocking (160).

(160) * It was shown that Jill seems that pigs fly.

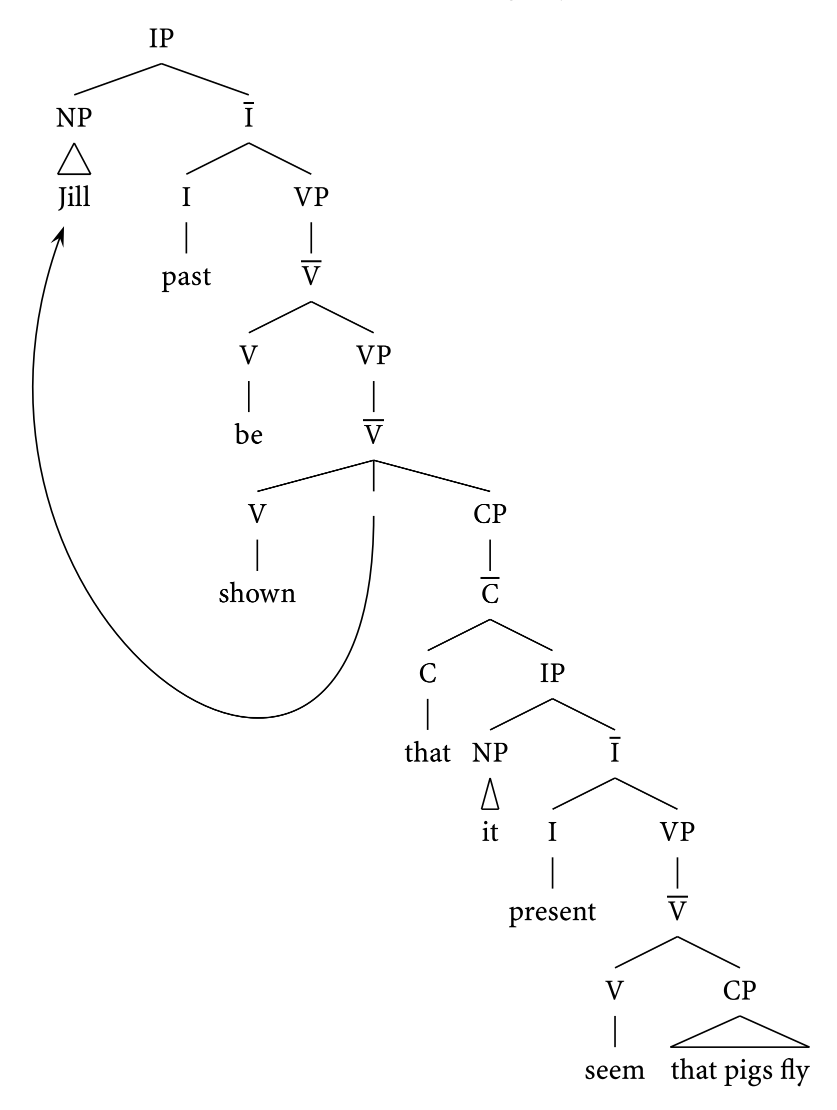
If Argument movement could move NPs to non-c-commanding positions, then (160) should be able to have the derivation indicated in (161) on the next page. The derivation satisfies every other condition we've discovered. The EPP is satisfied for both Specifier of IPs. The number of θ-roles and arguments is the same, so the Theta Criterion is satisfied. And every argument stands as a sister to the term that assigns it a θ-role at D-structure, and so the Projection Principle is satisfied as well. In particular, the argument *Jill* stands as a sister to the verb *shown* at D-structure, and there gets the object θ-role from *shown*. Because *show* has been Passivized, however, it does not have an accusative Case to assign to *Jill*, and so Argument movement must move this NP to a Case marked position. The Specifier of the finite IP that follows *Jill* is such a Case marked position, and so movement to this position should satisfy the Case filter. If this position c-commanded the D-structure position of *Jill*,

(161)



every other condition would be satisfied and this sentence should be grammatical. It's the c-command requirement on Argument movement, then, that is responsible for preventing sentences such as (160). This can be appreciated by comparing the grammatical (162), in which the c-command requirement is satisfied, with (160).


(162) Jill was shown that it seems that pigs fly.



If PRO is an argument NP, then we should expect that Argument movement will be responsible for fixing its position and, consequently, we should find a parallel pattern of grammatical and ungrammatical sentences that the successful or unsuccessful application of Argument movement is responsible for. In the case of overt NPs, the success of Argument movement hinges on being able to find an S-structure that satisfies the Case filter that can be produced from a grammatical D-structure through an application of Argument movement that satisfies the c-command requirement. In the case of PRO, the grammatical sen-

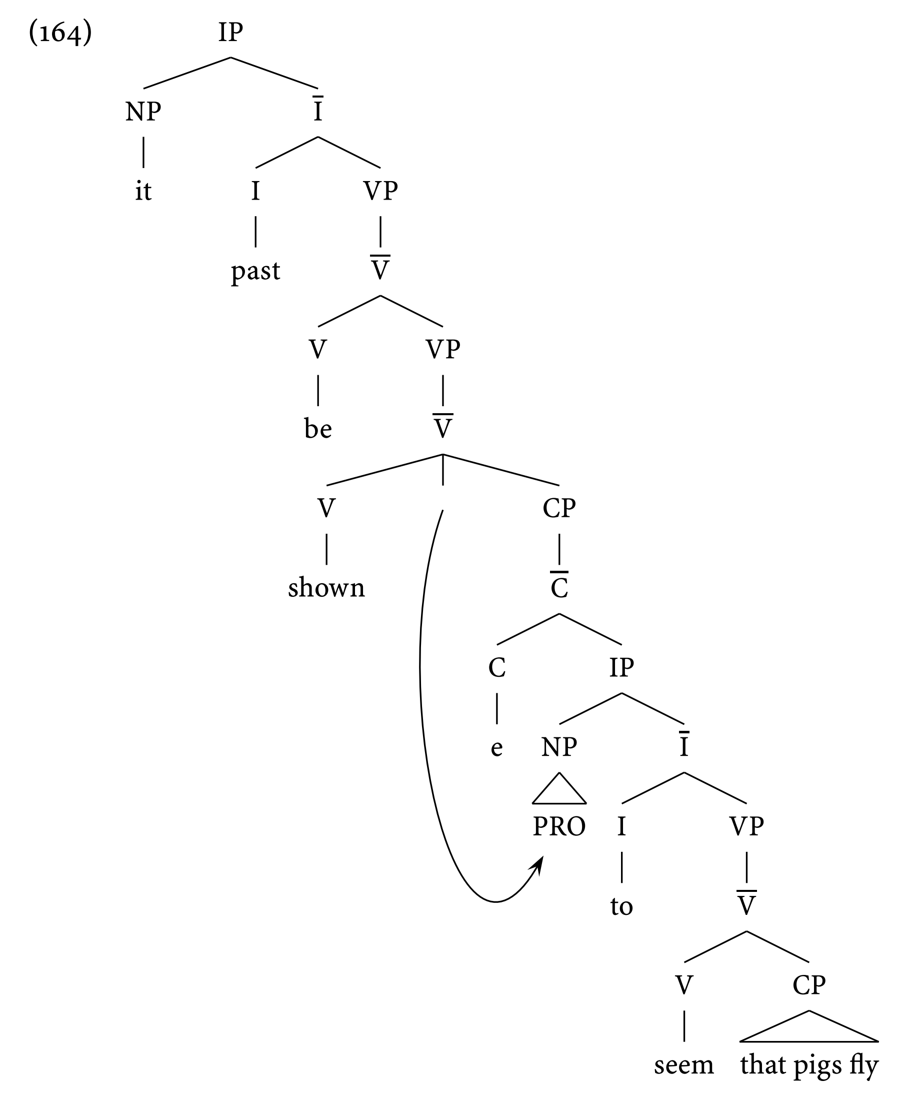
tences should be ones in which PRO can reach the Specifier of an appropriate infinitival IP from its θ-marked D-structure position by way of an application of Argument movement that satisfies the c-command requirement.

This match in grammaticality patterns shows up. The ungrammaticality of (163), for instance, parallels the ungrammaticality of (160).

(163) * It was shown to seem that pigs fly.

If the infinitival clause in (163) is of the sort in which PRO is permitted, then this example satisfies every condition except the c-command constraint on Argument movement, as (164) indicates.

(164)

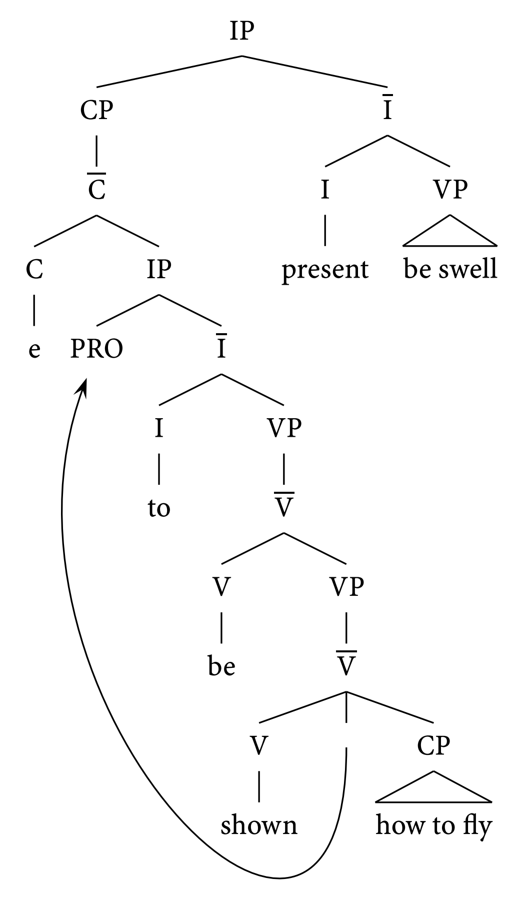


That this infinitive is of the kind that permits PRO in its Specifier position is indicated by (165), which permits an arbitrary reading for the subject of its infinitive.

(165) Jill was shown how to fly.

When the c-command requirement on Argument movement is obeyed, then examples parallel to (164) are grammatical, as (166) indicate.

(166) To be shown how to fly is swell.



Examples like these, then, suggest that the position of PRO is controlled by the same set of constraints that we've seen determine the position of overt argument NPs. To the extent that this is true, it speaks on behalf of the existence of PRO. If the ungrammaticality of sentences like (160) are due to conditions on where NP arguments may appear in S-structures, and those conditions are mimicked by arguments which are semantically present but not spoken, then

we should conclude that the syntactic position of those unspoken arguments matters. They must be present in the syntactic representation for that to hold.

There is another condition on the Argument movement of overt NPs that we can seek in the Argument movement of PRO. This condition we have not yet encountered; it concerns how the Case filter is satisfied. It appears that NPs cannot be Argument moved from positions that are Case marked.

(167) If Argument movement moves something from position α, then α cannot be a Case-marked position.

This condition is thought to be responsible for the ungrammaticality of (168).

(168) * Elena seems to that pigs fly.
(*compare:* It seems to Elena that pigs fly.)

The argument NP *Elena* resides in its θ-marked position, which is the complement position of *to*, and from there moves into the Case marked Specifier of IP position, thereby producing a representation that satisfies the EPP. The D-structure and S-structure representations meet the well-formedness requirements that they need to meet, and the position that *Elena* moves into c-commands the position it moves from. Even though it meets all the other requirements on a grammatical sentence, (168) is ungrammatical because it violates (167). The position that *Elena* resides in at D-structure is assigned accusative Case by *to*, and (167) prohibits Argument movement from a Case-marked position.

Another example whose ungrammaticality relies on (167) is (169).

(169) * Elena seems that is happy.

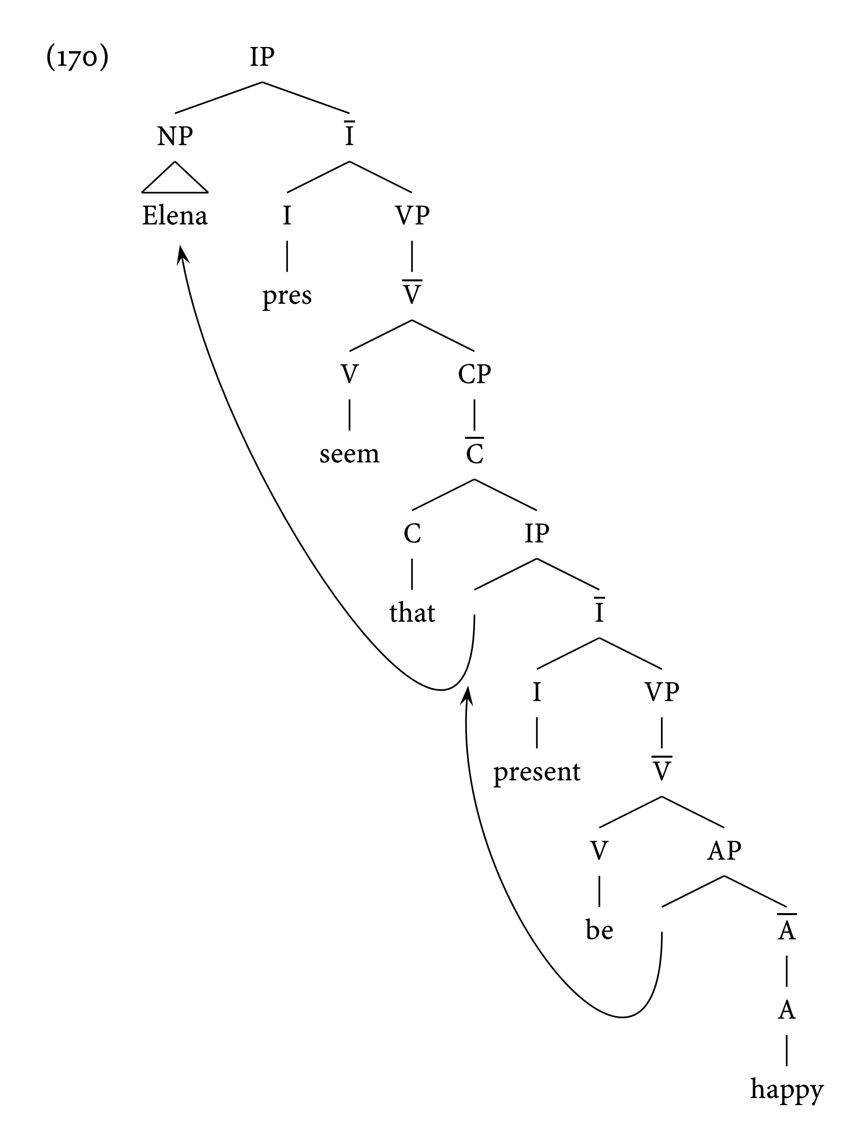
The derivation that needs to be blocked here is shown in (170) on the following page. For this sentence to satisfy the EPP, it is necessary for Argument movement to bring *Elena* into the Specifier of the embedded IP. But this position is assigned nominative Case by the finite I, and so movement from this position is prohibited by (167). There is consequently no way to derive (169) that does not violate either the EPP or (167).

There are examples parallel to these involving PRO, and they too speak on behalf of the existence of PRO and letting Argument movement play a role in positioning PRO. Parallel to (168) is (171).

(171) * To seem to that pigs fly is swell.

This example would be grammatical if it were possible to Argument move PRO from its D-structure position following *to* into the Specifier of the infinitival

(170)



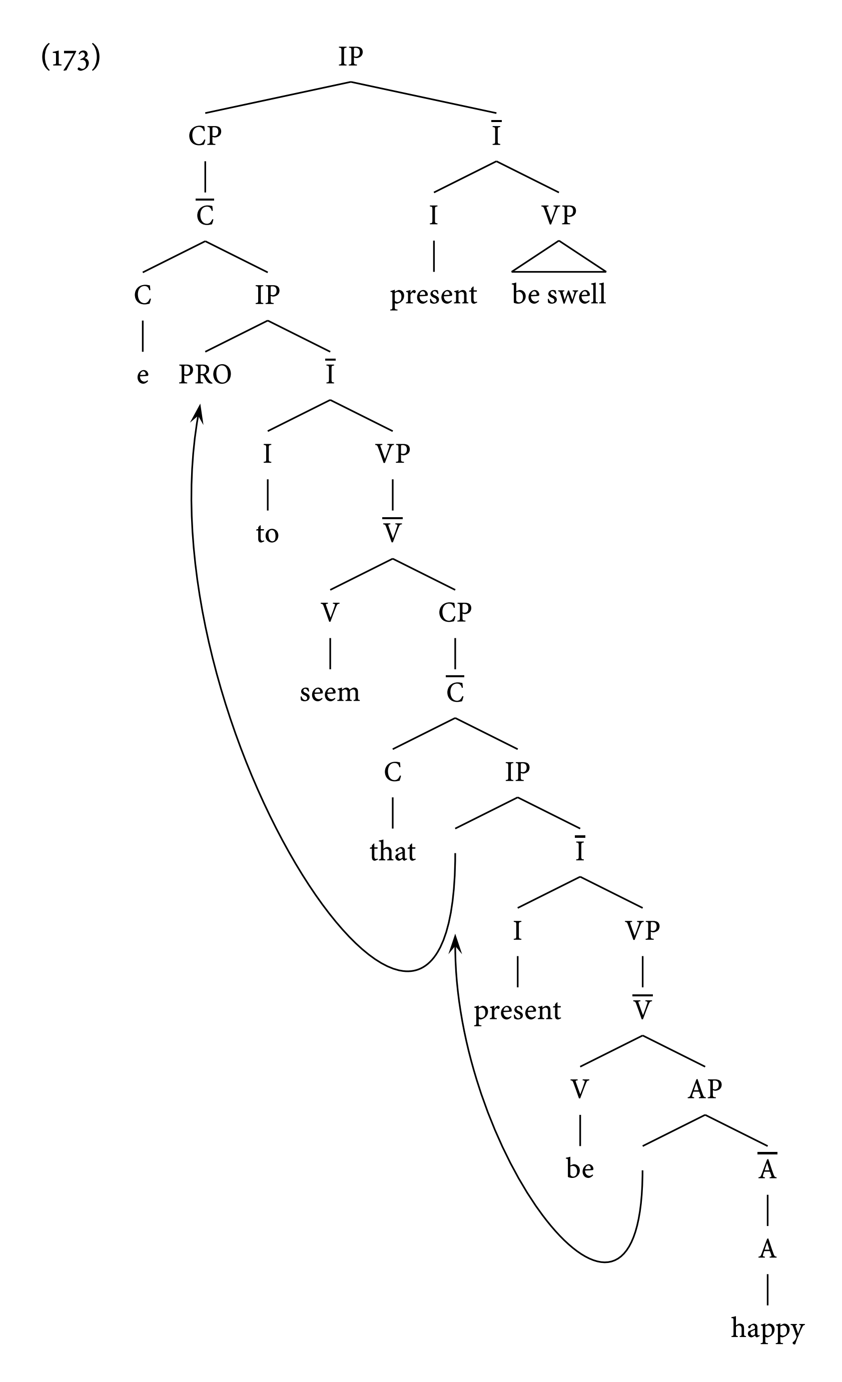
clause, where it will satisfy the EPP and also meet the condition on where PRO is allowed. To explain its ungrammaticality in a way that relates it to (168), we should credit the ban against Argument movement from Case marked positions.

Parallel to (169) is (172).

(172) * To seem that is happy is swell.

The derivation we should be considering for (172) — the one that satisfies the EPP — is (173).

(173)



Just like (170), this violates the ban against Argument movement from a Case marked position.

A final condition influencing Argument movement that might be relevant

for the placement of PRO is the Case filter. If PRO is an argument NP, then we should expect it to obey the Case filter. Is Argument movement invoked to put PRO into Case marked positions at S-structure just as it is to put overt NPs into Case marked positions?

It's not easy to determine the answer to this question because of the confound that the PRO Restriction presents. The PRO Restriction describes the fact that PRO seems to occur at S-structure in the Specifiers of control infinitives and nowhere else. If PRO is also subject to the Case filter, then the narrower distribution that the PRO Restriction allows it will make it difficult to detect. We are reduced, in fact, to the question of whether the positions that the PRO Restriction allows PRO to be in are Case marked positions.

We presently have two ways of determining whether a position is a Case-marked one. If an overt NP, bearing Case, is capable of surfacing in some position, then that position in assigned Case. And if Argument movement is prohibited from some position, then through a process of elimination we can determine whether the ban against moving from a Case-marked position is to blame and in this way discover whether a position is Case marked. By the first criterion, we cannot conclude the the Specifier positions in which PRO resides are Case-marked. The Specifiers of control infinitives cannot contain overt NPs.

(174) a. * Her to run seems unlikely.

b. * She to sing is swell.

From this, we might conclude that the positions PRO resides in are not Case marked positions. This would explain why overt NPs cannot surface in this position: the Case filter would be violated if so. This is the conclusion reached in *Lectures on Government and Binding*, and it informs most of the literature on PRO from the 1980s and 1990s.

The second criterion gives a somewhat more mixed image. If the movement account of Control is correct, then the ban against moving from a Case-marked position allows us to deduce that the Specifier of the control infinitive that a controller moves from is not Case marked. So, for instance, we can conclude that the Specifier position of the infinitival complement to a verb like *try* is not Case-marked if there is Argument movement of the sort indicated in (175).

(175) Hotze tried [$_{CP}$ [$_{IP}$ to eat durian]].

If the infinitive that controllers move from is the same kind that can host PRO in non-control environments, then the simplest assumption would be that the

Specifier of these infinitives is not Case-marked when they host PRO as well. So, for instance, if the infinitive clause that stands in the subject position of a Passivized *try* is the same as that which stands as the object of an Active *try*, as in (176), then the PRO this infinitive contains is standing in a non-Case-marked position.

(176) $[_{CP}$ $[_{IP}$ PRO to eat durian]] shouldn't be tried.

On the other hand, if the movement account of Control is incorrect, then we cannot use the prohibition against movement from a Case-marked position to deduce that the subject position of a control infinitive is Case-marked. And there are examples whose ungrammaticality could be derived from this prohibition if the subject position of control infinitives are Case-marked. One of these is (177).

(177) * Winnie was tried $[_{CP}$ $[_{IP}$ to eat durian]].

If the prohibition against Argument movement from a Case-marked position is not responsible for the ungrammaticality of (177), then it's not clear why it should be bad. There is only one θ-role for the argument *Winnie*, and so the Theta Criterion is satisfied, as is the Case filter and the c-command restriction on movement. Indeed, the contrast between (177) and (175) is problematic for the movement account of Control. The only relevant difference between these examples is the number of θ-roles involved: in (177), Passive has removed *try* of its external θ-role, whereas in (175) this θ-role remains. The movement account of Control relies on an as yet undiscovered force that will allow Argument movement from a control infinitive into a θ-marked position but prevent Argument movement from the same position that is not into a θ-marked position. A theory of Control that makes it a special reading for PRO, however, can make sense of this contrast with the tools at hand. Let the Specifier of a control infinitive be a Case-marked position, and everything follows. The Theta Criterion as well as the ban against movement from a Case-marked position will prevent Argument movement from a control infinitive into a θ-marked position. The movement that is illustrated in (175) will be blocked, and instead the subject of the embedded infinitival clause will be PRO. The ban against Argument movement from a Case-marked position will continue to prevent Argument movement from a control infinitive even when that movement is not to another θ-marked position, as in (177). On this set of assumptions, then, what makes a control infinitive different from, say, the infinitival complement

to a verb like *believe* is how the Specifier of the infinitive is Case-marked. The subject position of the infinitive that follows *believe* is accusative Case-marked by *believe*, and so Argument movement is possible from this position when *believe* is Passivized and thereby robbed of its accusative Case.

(178) a. Bart believed [$_{IP}$ him to be happy].
b. He was believed [$_{IP}$ to be happy].

The subject position of the control infinitive following *try*, by contrast, is Case marked by by something other than *try* and so Argument movement from that position will be blocked in both the Active and Passive versions of *try*.

(179) a. * Bart tried [$_{CP}$ [$_{IP}$ to be happy]].
b. * Bart was tried [$_{CP}$ [$_{IP}$ to be happy]].

If we adopt this account, the subject position of control infinitives is Case-marked and since these are the positions that PRO resides in, we can let PRO be subject to the Case filter.

What's mysterious under this account, though, is why overt NPs cannot stand in the subject position of control infinitives. If these positions are Case-marked, then we should expect not just PRO to be able to reside in them; overt NPs should be able to as well. But, as we've seen, this is not possible; examples like (180) are ungrammatical.

(180) a. * She to sing is swell.
b. * Elena tried her to sing.

A popular solution to this problem is proposed in Chomsky and Lasnik (1993). They suggest that the Specifier of a control infinitive is assigned a Case that only PRO may bear. English, they speculate, has a fourth Case, one whose morphological manifestation only PRO is capable of achieving. They call this the "null Case." This would have the desired effect of blocking overt NPs from the subject position of control infinitives. It could also be made to derive the PRO Restriction. If the only place that null Case is assigned is to the Specifiers of control IPs, and if we not only restrict null Case to PRO but prevent PRO from bearing any other Case, then the Case filter alone will properly segregate the positions of PRO and overt NPs. Overt NPs will have to surface in a nominative, accusative or genitive Case-marked position, and PRO will have to surface in a position marked with null Case. Martin (1992, 2001) argues that the availability of null Case can be tied to the semantics of the infinitival *to* that comes

with control infinitives. He argues that the *to* which heads control infinitives has an interpretation as a kind of tense, whereas the *to* that heads the infinitives that PRO cannot appear in does not have a tense interpretation. If correct, this would suggest that null Case is assigned under conditions rather like those that nominative Case is assigned. Nominative Case is assigned to the Specifier position of a finite I^0; on Martin's view, null Case is assigned to the Specifier of another type of tense I^0, the one that appears in control infinitives.

There are problems with both the idea that there is a special Case reserved for PRO and the view that this special Case is tied to a particular interpretation that distinguishes control infinitives from others. Pires (2001, 2006) and Terzi (1992) and Roussou (2001) (among others) have shown that the tensed interpretation Martin claims to correlate with the presence of PRO doesn't. Therefore, the supposed Case that only PRO can receive does not seem to be predictable from other properties of the infinitives it is alleged to be present in. Moreover, there is evidence that the subject position of control infinitives is capable of being assigned the normal Cases that can be borne by overt NPs. This evidence comes from Icelandic.

In Icelandic, floated quantifiers can be marked for Case and when they are, they have the same Case that the subjects they are related to do. In (181), for instance, the subject *Strákarnir* ('the boys') is inflected for nominative Case, and so is the floated quantifier *allir* ('all').

(181) Strákarnir komust allir í skóla.
the boys.nom got all.nom to school
'The boys all managed to get to school.'
(Sigurðsson 1991, (6a): 331)

The rules of Case assignment differ interestingly from English. Sometimes, the NPs in Specifiers of finite IPs can appear in something other than nominative Case. This seems to depend on the verb involved. For instance, the subject argument of the verb *vantaði* ('lacked') appears in the accusative Case, and the subject argument of *leiddist* ('bored') shows up in the dative Case. In these situations, a floated quantifier related to the subject will inflect for the Case that the subject bears. This is illustrated in (182).

(182) a. Strákana vantaði alla í skólann.
the boys.acc lacked all.acc in the school
'The boys were all absent from school.'

b. Strákunum leiddist öllum í skóla.
the boys.dat bored all.dat in school

'The boys were all bored in school.'

(Sigurðsson 1991, (6b,c): 331)

In Icelandic, then, we can tell what Case a subject NP has by looking at its floated quantifier; and subject arguments have a wider range of Case options than they do in English. We can use this feature of Icelandic to determine what Case, if any, PRO bears. What we find is that PRO seems able to bear whatever Case an overt NP would bear in that position.[14] The examples in (183) show this correspondence.

(183) Strákarnir vonast til ...
the boys.nom hope for ...
'the boys hope ...'

a. [$_{CP}$ að PRO komast allir í skóla].
[$_{CP}$ that PRO get all.nom to school]
'to all manage to get to school.'

b. [$_{CP}$ að PRO vanta ekki alla í skólann].
[$_{CP}$ that PRO lack not all.acc in school]
'to all not be absent from school.'

c. [$_{CP}$ að PRO leiðast ekki öllum í skóla].
[$_{CP}$ that PRO bore not all in school]
'to not all be bored in school.'

(Sigurðsson 1991, (8a–c): 331)

The controller — *Strákarnir* ('the boys') — is marked with nominative Case in each of these examples. But the quantifier floated in the control infinitive inflects in a way that depends on the verb of the infinitival clause. When the verb is the kind that makes its subject appear in the accusative or dative Case, in finite clauses, it makes the floated quantifier in the control infinitives inflect for the accusative or dative Case. If the match between the Case borne by a floated quantifier and the Case borne by the subject it is related to holds in infinitival clauses as well, then we can conclude from these examples that PRO bears the nominative, accusative or dative Cases assigned to Specifiers of finite IP in Icelandic.

Using null Case, then, to segregate PRO from overt NPs does not seem attractive. The evidence we have does not suggest that PRO and overt NPs differ with respect to the Cases they can bear. At present, it seems that the PRO

14 But PRO is capable of bearing other Cases as well, and so they do not behave precisely like overt NPs in this regard.

Restriction, and the ban against overt NPs in the subject position of control infinitives, must be sought elsewhere.

3.7.4 PRO and Control: A Summary

The evidence we've reviewed in this section, and the explanations of it that we've entertained, are complex and open-ended. It might be useful to arrange in one place some of the conclusions we've reached and hypotheses we've looked at.

We've discovered that there are silent argument NPs — "PRO" — which have a semantic interpretation rather like *one*. It shows up in infinitival clauses that function as subjects and as indirect questions; (184) are examples.

(184) a. PRO to eat durian is ill-advised.
b. I wondered how PRO to eat durian.

These infinitival clauses can be CPs, as the presence of the wh-phrase in (184b) shows. In closely related languages, like Icelandic, it's possible to show that they are always CPs. We've entertained the hypothesis that these infinitives are always CPs in English as well, and this requires the conclusion that the complementizers of these CPs are phonetically silent. Let's call these infinitives: "PRO-infinitives."

The existence of PRO means that there are forces which restrict its occurrence to just the places where we have evidence for it. This seems to be the Specifier of the IPs in PRO-infinitives. These same positions do not seem to allow overt NPs, as the contrast between (185) and (184) indicates.

(185) a. * She to eat durian is ill-advised.
b. * I wondered how him to eat durian.

We have entertained the hypothesis that what makes the Specifier positions of PRO-infinitives special in this way has to do with the Case they are assigned. But so far as I can tell, there is no evidence that these Specifiers get assigned a special Case, one that would make it different from the Cases assigned to other Specifiers of IPs. I suggest we abandon this hypothesis, therefore.

So, putting these parts together, we have:

(186) a. There is a silent NP: PRO.
b. PRO can get a semantic interpretation like that assigned to *one*: this is called its "arbitrary" interpretation.

c. PRO can only appear in the Specifiers of PRO-infinitives.
d. Overt NPs cannot appear in the Specifiers of PRO-infinitives.
e. PRO-infinitives can be CPs. It's possible that they are always CPs.

We've also encountered infinitival clauses that do not contain a PRO with an arbitrary reading, but instead have a silent subject that refers to the same thing that an argument spoken elsewhere does; (187) is an example.

(187) Andre tried to sing.

In these cases, the overt NP 'controls' the understood subject of the infinitive. We have two ideas about how to characterize these situations.

On the first view, control is a special instance of Argument movement, and so the syntax of (187) is identical to that posited for (188).

(188) Andre seemed to sing.

In both cases, *Andre* moves from the position *sing* assigns a θ-role to. In (188) this is the only θ-role *Andre* receives, as the verb *seem* does not assign an external θ-role. In (187), by contrast, *Andre* also receives a θ-role from *try*, presumably by moving into the position this θ-role is assigned to. If this is the correct treatment of Control, then our present formulation of the Theta Criterion will have to be changed so that it allows on argument to bear more than one θ-role. This account of Control will also have to find a way of capturing the difference between (187) and (189).

(189) * Andre was tried to sing.

And, if the Specifiers of these infinitives are Case-marked, as the evidence from Icelandic suggests, then (187) would seem to be a violation of the prohibition against Argument movement from Case-marked positions. So, this characterization of Control includes the points in (190).

(190) If Control is Argument movement, then:
a. PRO-infinitives are prevented from arising in contexts where Control is obligatory.
b. The Theta Criterion requires changing.
c. There is an unknown force responsible for the contrast in:
i. Andre tried to sing.
ii. * Andre was tried to sing.

d. If the Specifiers of the IPs from which Control arguments move is Case marked, then the generalization that Argument movement from Case-marked position is ungrammatical is lifted in these contexts.

On the second view, Control is a special meaning that PRO is able to receive. If this is correct, those infinitives in which Control arises are PRO-infinitives. This characterization of Control brings with it the following points.

(191) If Control is found in PRO-infinitives, then:

a. PRO cannot get an arbitrary reading in those contexts where Control is obligatory.
b. There is a semantic rule independent of Argument movement which gives the Control reading.
c. If the Specifiers of PRO-infinitives are Case-marked, then the ban against Argument movement from Case-marked positions can be used to account for the contrast between:
   i. Andre tried to sing.
   ii. * Andre was tried to sing.

These two approaches to the Control relation are presently hotly contested. I will try to stay neutral between the two views and their consequences in what follows. But if it becomes necessary to choose, I will pick the second view. If that should happen, I'll also tell you why.

3.8 Argument Movement and a typology of verbs

One can think of Argument Movement as a function from a parse that satisfies the Projection Principle and Theta Criterion to a parse that satisfies the Case Filter. θ-roles are determined by lexical items and the $\overline{\text{X}}$s they project, and where Case is assigned is also controlled by the these lexical items. Lexical items, and verbs in particular, play a big role in setting up the environments in which Argument movement is called upon. The conditions on Argument movement place a cap, then, on the kinds of situations that verbs can create. They won't allow verbs to exist which create conflicts which Argument movement can't fix.

Verbs set up these situations by influencing where the θ-marked and Case marked positions for an argument are. Thus, for example, we have seen that

there are verbs which assign one external θ-role, and verbs which assign an external and an internal θ-role and Accusative Case. These are sometimes called intransitive and transitive verbs, respectively; examples are in (192).

(192) a. Sally slept. (intransitive)
b. Sally likes kiwis. (transitive)

A special kind of transitive verb are exemplified by *believe*, *consider* and *make*, which assign their Accusative Case to something different than the argument they assign an internal θ-role to. As we've seen, these verbs can take a "clausal" complement — sometimes these are "small clauses" — and assign their Accusative Case to an NP within this clause.

(193) a. She believes [$_{IP}$ him to be unhappy].
b. She considers [$_{AP}$ him happy].
c. She made [$_{VP}$ him dance].

We have also seen verbs that have no external θ-role but do have an internal θ-role. One of these we have seen in situations such as:

(194) Sally appears [$_{IP}$ to like kiwis].
Sally seems [$_{AP}$ happy].

And others we have seen formed by the process of passivization:

(195) a. Sally was considered [$_{AP}$ unhappy].
b. Sally was considered .

There is evidence that some verbs which might otherwise look like intransitive verbs fall into this last class too. For instance, when *appear* c-selects not a clause, as it does in (194), but an NP, this NP surfaces in the nominative Case marked subject position, as shown in (196).

(196) A ghost appeared.

Burzio (1986), who produces the first systematic arguments on behalf of these two classes of single argument verbs, uses the terms *unergative* and *ergative* to distinguish them. Others, notably David Perlmutter who is the co-discoverer of this distinction,[15] have used the term *unaccusative* for what Burzio calls ergative verbs. Let's use the term *intransitive* as a label for either type of single-argument verb, with these two sub-classifications.

15 See Perlmutter (1978)

(197) Intransitives

a. A ghost should sleep. (unergative)

b. A ghost should appear. (ergative, aka unaccusative)

There are a wide assortment of syntactic phenomena that are sensitive to the distinction between these two sorts of intransitive verbs. We will encounter a few of them in the lectures that follow. In English, one of the phenomena that confirms the picture that there are intransitives that have a single internal argument as well as those that have a single external argument comes from the adjective formation process we briefly discussed in the previous chapter. Recall that this rule creates from a verb an adjective whose external θ-role is the same as that assigned by the verb to its "direct" object. This process, then, should only be able to apply to verbs that have a direct object θ-role, and indeed it is blocked for a large range of intransitive verbs as a consequence.

(198) a. * the danced man

b. * the ran dog

c. * the slept woman

d. * the cried child

But there are a small class of intransitives which are able to be adjectives by this process, and these seem to be just the ergative or unaccusatives.

(199) a. the fallen leaves

b. the recently arrived package

In other languages, there is a wider assortment of phenomena that appear to be sensitive to the unaccusative/unergative distinction in intransitives.

Let's consider, then, the space of verb types that we might expect to find and compare that to what we have found so far.

(200)

θ-roles	Accusative Case	No Accusative Case
External, no internal	??	unergatives (*sleep*)
No external, internal	??	ergatives (*appear*)
external and internal	transitives (*like*)	??
no external, no internal	??	??

If assigning an external θ-role, an internal θ-role and assigning Accusative Case are independent properties that verbs have, then this table expresses all the possible ways in which we might expect these properties to combine. As

can be seen, there are quite a number of verb types that we haven't yet seen, but which we should expect to exist.

Are there verbs yet to be discovered that fill these various cells? In some cases, the properties combined make for verbs that are rather difficult to discover. Consider, for instance, a verb that assigns an external θ-role and Accusative Case, but assigns no internal θ-role (a verb that would fit in the cell in the top row, first column). It will be very difficult to discover verbs of this sort, even if they should exist, because without a complement there will be nothing to bear the Accusative Case that such a verb would assign. The only way to see such a verb would be in cases where we might find a non-complement to which, or into which, Accusative Case could be assigned. One candidate, perhaps, for this situation are cases such as (201).[16]

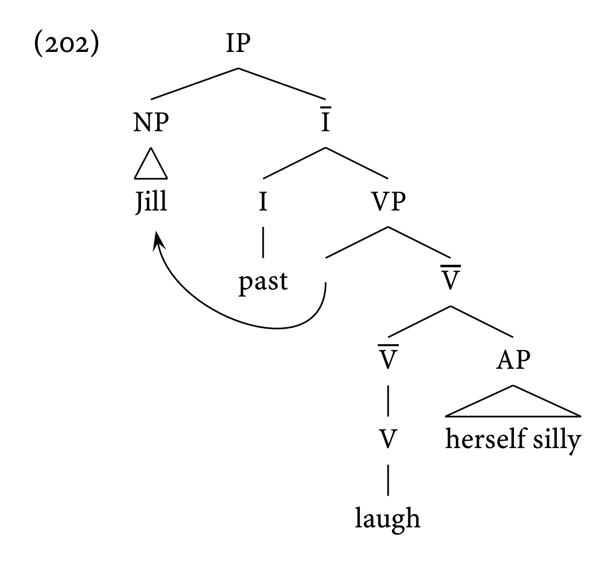
(201) Jill laughed herself silly.

It is likely that the small clause, *herself silly*, is not a complement to *laugh*; it does not, in any case, refer to something that is involved in the event, or action, that *laugh* denotes. If the Accusative Case on *herself* comes from *laugh* — and where else could it come from? — then *laugh* is a verb of the sort we are in search of.

It should be noted that this analysis of (201) is at odds with some of the rest of the grammar we have developed. If *herself silly* is not θ-marked by *laugh*, then the Projection Principle is going to require it to be a sister to the $\overline{\text{V}}$ that *laugh* projects, as indicated in (202).

16 My thanks to Angelika Kratzer for suggesting that I use the adjective formation rule as a diagnostic for unaccusatives, and for offering this construction as an example of this class of verb. See Carrier and Randall (1992) and Kratzer (2005) for some discussion of this construction.

(202)



But in this configuration, *laugh* will not c-command *herself*, and this is a requirement on Accusative Case assignment. Indeed, if the c-command requirement on Case assignment is correct and the Projection Principle's placement of complements is too, then these will conspire to prevent verbs of the sort we are searching from ever being found. If *laugh* genuinely is such a verb, then these parts of our grammar will need adjustment. This is work for the future.

Consider now verbs that assign neither an external nor an internal θ-role: the class of verbs that would fill the cells of the bottom row in (200). Do these verbs exist? A candidate are verbs such as *rain*.

(203) It rains.

If the *it* in this example is not an argument, then here is a verb that assigns no θ-role.

What of the other two categories of missing verb? Are there verbs which support no external θ-role, but do assign an internal θ-role and Accusative Case? And are there verbs that assign both external and internal θ-roles, but no Accusative Case? To date, there are no verbs with these properties that have been discovered in English. At present, then, we can update the table in (200) to (204).

(204)

θ-roles	Accusative Case	No Accusative Case
External, no internal	*laugh*?	unergatives (*sleep*)
No external, internal	not found	ergatives (*appear*)
external and internal	transitives (*like*)	not found
no external, no internal	??	*rain*?

Burzio discovered the two gaps in this paradigm where there appear to be no verbs, and formulated generalizations which describe these absences.

(205) Burzio's Generalization

a. If a verb assigns accusative Case, then it assigns an external θ-role.

b. If a verb assigns an external θ-role (and an internal θ-role?), then it assigns accusative Case.

Why haven't we found verbs like these? Burzio's generalizations might reflect a relationship between Accusative Case and θ-role assignment for which we should find a source.

The second of these generalizations might be derivable from the conditions we have seen on Argument movement. To see this, consider the syntactic frames that our theory would let this verb be inserted into. One of these is (206).

(206) $[_{\text{IP}}$ should $[_{\text{VP}}$ Smith V Jones]].

If V assigns these two θ-roles, but no Case to *Jones*, then there is no way both of these NPs are going to be able to satisfy the Case Filter. There are more NP arguments than there are Cases. So, if such a verb is to survive, the only environment it will be able to appear in are sentences which have two Case marked positions. These two positions will both have to be Specifiers, because these are the only positions reachable by Argument movement. Thus, we're looking for contexts like:

(207) $[_{\text{IP}}$ – should V_1 $[_{\text{XP}}$ – $[_{\text{VP}}$ Smith V_2 Jones]],
where both "–" positions are Case marked.

Now the first part of Burzio's Generalization tells us that V_1 cannot assign Accusative Case. If it did, then it would also assign an external θ-role, and that's going to bring the count of things that need Case to one more than there are Case marked positions. As a consequence, the Case marked position inside XP is going to also have to get its Case from some place other than V_1. So far, the only ways we have seen for this to be done are if XP is in a CP:

(208) $[_{\text{IP}}$ – should V_1 $[_{\text{CP}}$ for $[_{\text{IP}}$ – to $[_{\text{VP}}$ Smith V_2 Jones]]]].
$[_{\text{IP}}$ – should V_1 $[_{\text{CP}}$ that $[_{\text{IP}}$ – I^0 $[_{\text{VP}}$ Smith V_2 Jones]]]]
$[_{\text{IP}}$ – should V_1 $[_{\text{CP}}$ e $[_{\text{IP}}$ – to $[_{\text{VP}}$ PRO V_2 Jones]]]]

The only place a verb of this sort could find a grammatical outcome is when it is embedded under a verb that does not assign an external θ-role. That is arguably too narrow a niche for such a verb to become salient enough to be acquired. Moreover, we will see in a later chapter that there is some evidence for a constraint on Argument movement that prevents one argument from moving past another c-commanding argument. If such a constraint does exist, it would prevent *Jones* from reaching either of the Case-marked positions in (208). Consequently, even this one environment for such a verb would be removed.

For this reason, the usual conclusion is that the second of Burzio's generalizations derives from constraints on Argument movement in concert with the Case filter. What remains is (209).

(209) Burzio's Generalization
If a verb assigns accusative Case, then it assigns an external θ-role.

If Burzio's Generalization cannot be derived from constraints on the syntactic form that such verbs would have to live in, then it suggests that the assignment of Case and θ-roles is not independent in the way that we presently think it is. We will need to return to this issue.

But before we do that, let's turn to some of the other matters we left unfinished in the preceding chapter. In this chapter we've managed to clarifying enough of the forces that seem responsible for giving arguments their syntactic positions to tie up many of those loose ends. We now know what lies in the Specifier positions of many phrases and, at the same time, we have an idea about where subject θ-roles are assigned. We have developed an account of the conditions that give a linear order to complements which is based on Case assignment, and we've coupled this with a theory that allows arguments to be moved. This movement theory claims that one movement rule — Heavy NP Shift — is partly responsible for giving complements their post-verbal position, and another movement rule — Argument movement — is responsible for given NPs their S-structure position when they are not θ-marked in Case-marked positions. Along the way, we've discovered evidence for the existence of a silent argument NP, and this has generated some additional open questions. We've answered, then, all of the questions we set out at the beginning of this chapter to answer, except for one. We have left the question of how the inflectional morphology that our grammar places in I^0 manages to get expressed as part of a verb that follows, rather than as a morpheme in the position our rules place it. Before we return to the questions we've opened in this chapter — questions about how to control the distribution and interpretation of PRO, and

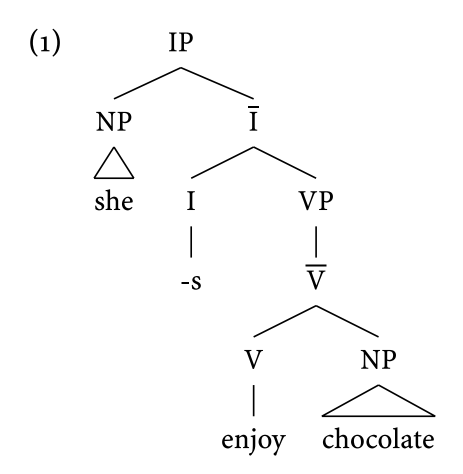
the connection between θ-role assignment and Case that Burzio's Generalization seems to point to — let's clear up this matter about I^0. Surprisingly, what we learn about the expression of inflectional morphology in I^0 lends a hand to these other questions.

4

Verb Placement and Features

One problem with our attempt to characterize the grammatical English sentences has to do with the fact that the heads of sentences are often not words. Many finite clauses are headed by bound morphology, as in (1).

(1)



We reached this conclusion because it is morphology of this sort — agreement/tense morphology — that is in the one-to-one relation with the family of strings that we call IP. What we need to understand is how it is that this bound morpheme manages to find itself expressed on the verb that follows.

There is some evidence that in certain cases of this sort, the verb moves to the position our rules put the inflectional morphology. The rule responsible is, therefore, sometimes called Verb Movement. Its classic description is found

in Chomsky (1957) and Emonds (1976), and it is this process that is examined in our reading: Pollock (1989).

4.1 The Classic account

Chomsky and Emonds use a type of reasoning that I will dub a "Correlation Argument," because it's based on the correlation that holds between a verb's syntactic position and its inflection. The generalization that accounts like theirs hope to derive is the dependence between a verb's syntactic position and its inflectional class. This dependence can be seen from the relative positions of auxiliary verbs in English finite clauses. An auxiliary verb cannot follow *not, so*, *too* ("polarity terms") if it is inflected with the tense/subject-agreement morpheme. We find paradigms like those in (2).

(2) a. Andre has not been eating.
   b. Andre is not eating.
   c. * Andre not has been eating.
   d. * Andre not is eating.
   e. Andre has too/so been eating.
   f. Andre is too/so eating.
   g. * Andre too/so has been eating.
   h. * Andre too/so is eating.

But when an auxiliary verb is not inflected with tense/subject-agreement morphology it can follow these polarity terms.

(3) a. Andre must not have eaten.
   b. Andre must not be eating.
   c. Andre will too/so have eaten.
   d. Andre will too/so be eating.

In these contexts we find that the polarity item must be found immediately before the auxiliary verb.

(4) a. * Andre not must have eaten.
   b. * Andre not must be eating.
   c. * Andre must have not eaten.
   d. * Andre must be not eating.

e. * Andre too/so will have eaten.
f. * Andre too/so will be eating.
g. * Andre will have too/so eaten.
h. * Andre will be too/so eating.

Some English speakers may judge (2c,d) grammatical. For these speakers, I suspect that *not* is being used differently than it is in situations where *not* appears between the modal and auxiliary, as in (5).

(5) a. Andre will not have eaten.
b. Andre will not be eating.

In (5), *not* negates the entire sentence. The speaker of these sentences means to deny the un-negated version of these sentences, i.e. the sentences in (6).

(6) a. Andre will have eaten.
b. Andre will be eating.

By contrast, the sentences in (2c) and (2d) are positive assertions. The speaker of these sentences means to assert that Andre must have done something – namely not eaten – or must be doing something – namely not eating. This difference between (2c,d) and (5) can be brought out by considering how they combine with so-called question "tags." These tags, illustrated in (7), have the opposite polarity as the clause they are appended to.

(7) a. Andre will eat, won't he?
b. Andre will not eat, will he?

If the polarity of the tag is the same as the clause it's appended to, the result has a rhetorical flavor:

(8) a. Andre will eat, will he?
b. Andre will not eat, won't he?

When *not* stands between the modal and a following auxiliary, adding a positive question tag gives a question of the sort illustrated by (7); cf. (9).

(9) Andre will not have eaten, will he (have)?

But when *not* follows both modal and auxiliary, adding a positive question tag yields the rhetorical sorts of questions in (8); cf. (10).

(10) Andre will have not eaten, will he (have)?

On the other hand, when the auxiliary is finite, and so in the I^0 position, a *not* that follows it behaves the same as the one that follows a modal:

(11) Andre has not eaten, has he?

To the extent that there are real judgments here – it's not easy to be clear about what the difference between the "real" and "rhetorical" questions is – it suggests that we do not have the same use of *not* when it stands before and after a non-finite auxiliary. If these uses of *not* correspond to the syntactic position it occupies, then we want to conclude that *not* is in the same position in (12), but a different position in (13).

(12) a. Andre will not have eaten.
 b. Andre has not eaten.

(13) Andre will have not eaten.

These data, then, lead us to conclude that polarity items lie between I^0 and VP. And from this we can conclude from the contrast between (2) and (3) that verbs are in I^0 when they bear the inflection that resides there. There is, in other words, a correlation between a verb's inflectional class and its syntactic position.

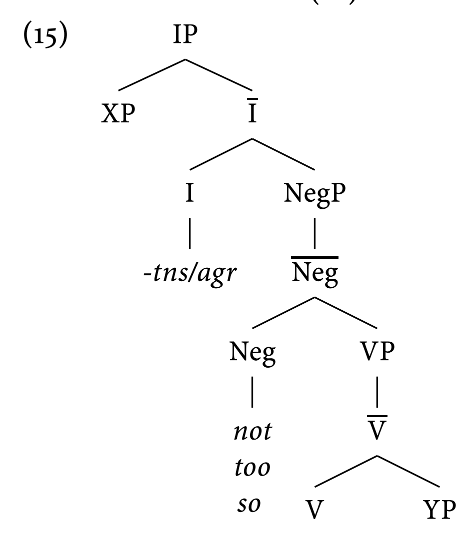
That auxiliary verbs have a different position when they are inflected is confirmed by the action of rules that affect VP. These rules seem to necessarily strand the verb inflected for agr/tense. This is indicated by the following examples of VP deletion and VP topicalization.

(14) a. Sam is eating pickles because Mike is △.
 b. Sam should be eating pickles because Mike should △.
 c. * Sam is eating pickles because Mike △.
 d. I claimed that Mary is eating pickles, and [$_{VP}$ eating pickles] she is.
 e. I claimed that Mary has to be eating pickles and [$_{VP}$ be eating pickles] she has to.
 f. * I claimed that Mary is eating pickles, and [is eating pickles], she.

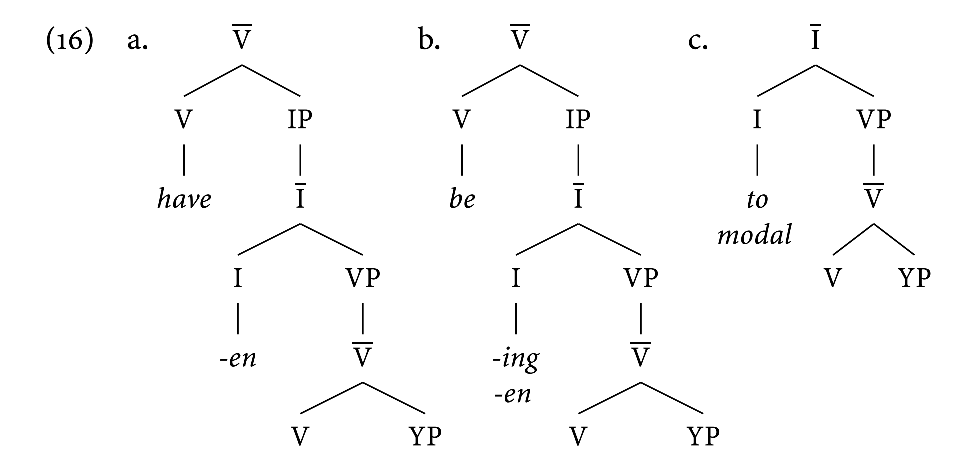
If VP Ellipsis is capable of eliding any VP, then the ungrammaticality of (14c) indicates that the finite form of *be* is not within a VP. Similarly, (14f) indicates that the movement rule responsible for moving a VP at the beginning of a clause cannot relocate a finite form of *be*. This would follow if finite forms of *be* are not within the VP that is being moved in these constructions.

A Verb Movement account of this phenomena lets phrase structure rules fix the location of the inflectional morphemes involved, and moves the appropriate verb to this location by way of a transformational rule. Thus, simple finite clauses are assigned a D-Structure representation like that in (15) below, and clauses with a complex tense (i.e., an auxiliary verb), have D-Structure representations like those in (16).

(15)



(16) a. b. c.



In these parses I've decided to position polarity items as heads that embed the following VP. Moreover, I've renamed these items, and the phrase they head, "Negation." This renaming is common in the literature, as it is typical that only the negative item among the polarity terms is talked about. The representation in which Neg^0 embeds VP is argued for in Pollock (1989), and we will consider his reasons soon. There are other possibilities which are consistent with what

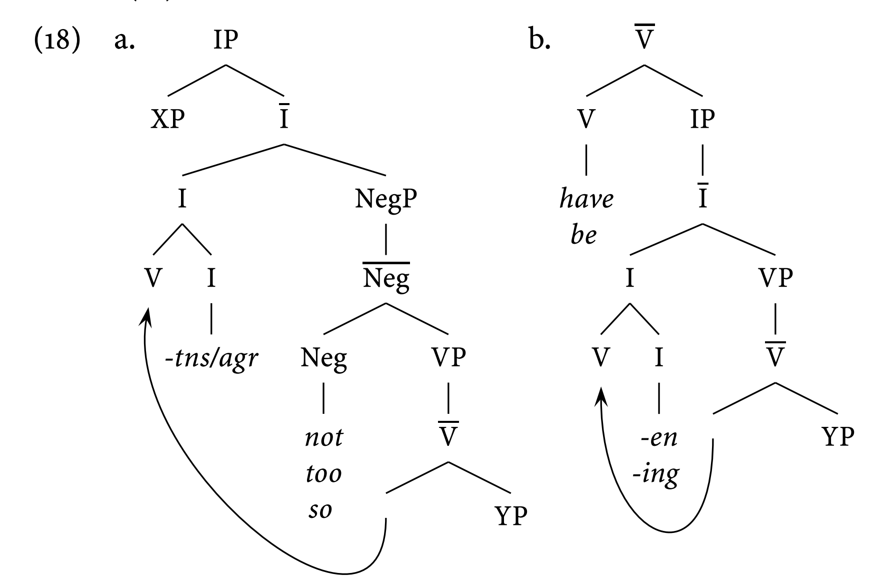
we've seen so far, however. One is to treat polarity items as heading a phrase that adjoins to the left side of $\overline{\text{V}}$, like a modifier would.

The Verb Movement rule could be formulated as in (17).

(17) Verb Movement
Adjoin V^0 to the left of I^0.

This will produce from the D-structures in (15) and (16) representations like those in (18).

(18)



This gives a straightforward account of the correlation described above. Verbs stand in the positions determined by the inflections they bear because they are moved to the positions that the phrase structure rules assign to these affixes.

To force Verb Movement in these contexts, we could posit a requirement that an affix must share an X^0 with its stem at S-structure. Lasnik (1981), where this idea is proposed, calls this condition the "stray affix filter."

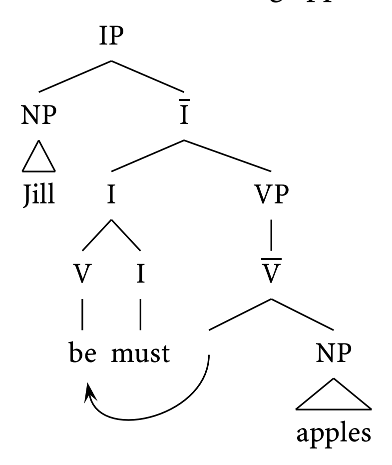
(19) Stray Affix Filter
An affix must co-exist with its stem under a common X^0 to be pronounced.

To prevent Verb Movement in contexts where the I^0 contains a word, such as a modal, we could strengthen the Stray Affix Filter to (20).

(20) The Word Criterion
Let α be an X^0 immediately dominated by $\overline{\text{X}}$. Everything α dominates must form one word.

The Word Criterion, then, will force Verb Movement when I^0 contains something less than a full word, and block Verb Movement in cases such as (21) on the facing page, where the result would put into I^0 more than one word.

(21) * Jill be must eating apples.



An exception to this scheme is encountered when English "main" verbs are considered. In these situations, the verbs do not occupy the positions determined by the inflections they bear, but remain instead in the location given to main verbs by the phrase structure rules. Examples like (22) are ungrammatical.

(22) *Andre likes not/too/so apples.

And the processes that indicate whether a verb is in its VP, i.e., VP ellipsis and VP fronting, indicate that finite main verbs are indeed positioned within the VPs they head.

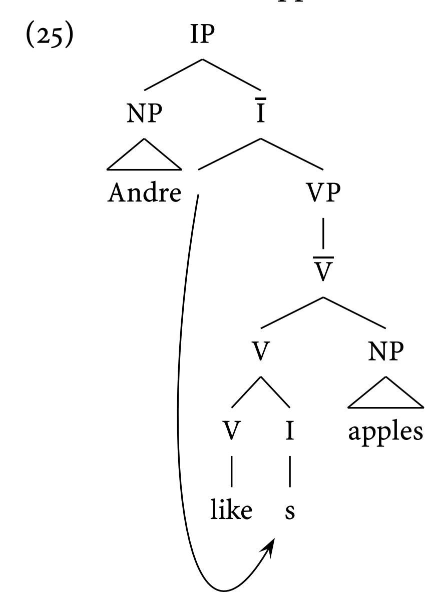
(23) a. * Andre ate apples because Jill △.
b. * I said that Andre ate apples and [$_{VP}$ ate apples] Andre.

The process that brings inflection together with the verb on which it is expressed must have a different outcome when main verbs are involved. The challenge is to find a way of conceiving of this process so that both kinds of outcomes are possible. There have been a number of attempts to capture this flexibility, none of them entirely successful.

One account, found in Chomsky (1957), posits a transformation rule that moves the inflection down onto the verbal root when main verbs are involved. A sentence like (24), for example, would enter into the derivation in (25). On


(24) Andre likes apples.

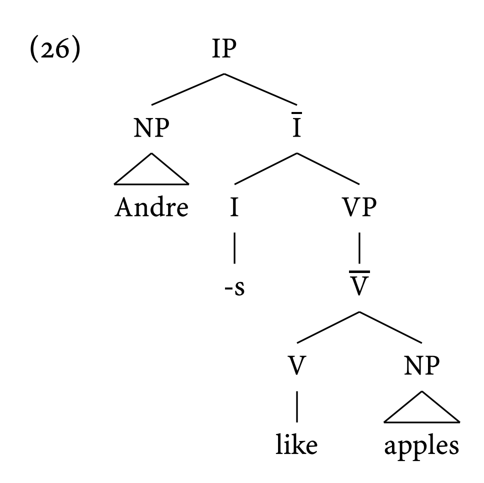
(25)



this view, then, verbs and their inflections always combine by way of a movement rule. The difference between auxiliary and main verbs concerns what moves: the verb, in cases involving an auxiliary, and the inflection, in all other cases. A problem with this proposal is that it requires movement transformations to be able to move things to non-c-commanding positions. We have seen that for other movement transformations, this does not seem to be possible. Outside of these contexts where main verbs get inflected without moving, the same generalization holds with respect to the rule that moves verbs around. It looks like it is worth preserving the generalization that movement rules do not move things to non-c-commanding positions. For the reason, the proposal in Chomsky (1957) has been largely abandoned.

A second account lets the mechanism that brings a main verb together with its inflection be a modification of how lexical insertion works. Presently, our formulation of lexical insertion matches lexical items to X^0s, and this prevents one word from spanning more than one terminal position in a phrase marker. Bobaljik (1994, 1995) argues that this should be weakened so that a lexical item can be inserted into a position that does not map directly onto a terminal in a phrase marker. In particular, he proposes to let an inflected form of a main verb be inserted into a position occupied by two X^0s, just in case those two X^0s are linearly adjacent. Under this proposal, a representation like (26) below could be directly mapped onto the string *Andre likes apples*.

(26)



This proposal has the benefit of explaining why finite sentences containing polarity items in English are ungrammatical if they do not contain an auxiliary verb.

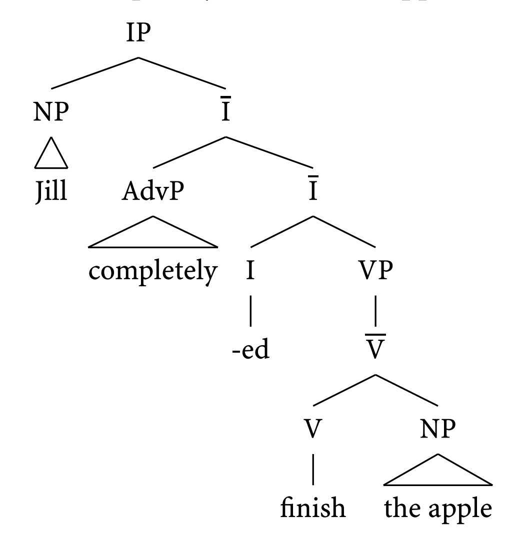
(27) a. * Andre not likes apples.
b. * Jill not ran down the street.
c. * Winnie so read that encyclopedia.
d. * Bart too ran into the street.

Because polarity terms are positioned between I^0 and the verb that follows, it will prevent an inflectional morpheme residing in I^0 from being adjacent to that verb and block lexical insertion from rendering them as a single inflected verb. Indeed, this proposal predicts that a finite clause with no auxiliary verbs in it should not allow any material to stand between I^0 and the main verb. The only grammatical parse for an example like (28) on the following page should be one in which the adverb is to the left of I^0, as indicated. This is not an obviously correct prediction, however. Contrasts like those in (29) indicate that *completely* must be placed between I^0 and the following VP.

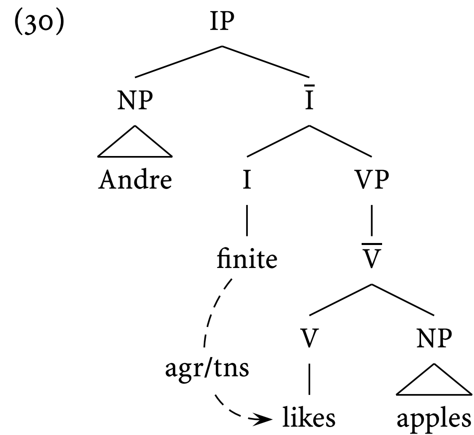
(29) a. Jill will completely finish the apple.
b. * Jill completely will finish the apple.

This account, then, has difficulties making sense of data like (28). Unlike Chomsky's proposal, it does address the problem of why the mechanism that inflects main verbs is sensitive to the presence of polarity items; but it doesn't explain how these items differ from other material that appears to be able to stand between the main verb and finite inflection.

(28) Jill completely finished the apple.



(30)



The final idea we will consider divorces the mechanism that brings a verb and its inflection together from the process that moves verbs around. On this view, the relationship between an I^0 that is associated with inflectional morphology and the verb that expresses that morphology is taken to be akin to the relationship we have posited between an NP and the term that assigns that NP Case. We can think of I^0 as assigning the relevant inflectional morphology to the verb that follows, as in (88) on page 195. Chomsky (1995), where this proposal can be found, calls the assignment relationship between an inflected verb and the I^0 where that inflection is determined: AGREE. The precise for-

mulation of AGREE will be taken up later; but it should be clear from (88) that it doesn't match perfectly the condition we arrived at for Case assignment in the previous chapter. While the finite I^0 does c-command the verb it assigns tense/agreement inflection to in (88), it is not klose to that verb: there is more than one phrase that dominates the verb but not I^0. This would block Case assignment. One desirable consequence of the formulation of AGREE is to make it account for the inability of a polarity item to stand between I^0 and a main verb. That would be a natural place for this account to attempt an explanation for this fact.

If finite I^0 assigns inflectional morphology to a main verb that follows, then there are a couple ways of characterizing the different outcome with auxiliary verbs. One possibility is that auxiliary verbs are inflected in the same way, and that there is an independent reason that they must move to the I^0 that assigns them their inflection. Another possibility is that auxiliary verbs are prevented from getting their inflectional morphology by way of AGREE, and instead make recourse to Verb Movement. Hopefully, a better understanding of what makes an auxiliary verb different from a main verb will help us understand how to express the difference in how their inflection influences their surface position. We can learn something about this question by considering how these processes arise in other languages. We'll look at other languages in the sections that follow, and then return to the question of how to give a picture of the main verb/auxiliary verb distinction in English. But first, there is one last piece to the picture of English verb placement that we should see.

4.2 Movement to C^0

The rules we've looked at that relate verbs to I^0 can be embellished with another that plays a role in forming questions. One sort of question involves relocating the finite verb to the beginning of the sentence, as in (31).

(31) a. Have you eaten pickles?
 b. Should you eat pickles?

These are called Yes/No questions. Another sort of question which involves a relocation of the finite verb seeks more specific information and is called a Wh-Question. In these questions, as we saw in the previous chapter, there is also movement of a phrase into Specifier of CP. Some examples are in (32).

(32) a. Which pickles have you eaten?

b. What should you eat?

In both of these types of questions, the contents of I^0 has moved to some sentence-initial position. As a consequence of this rule's targeting I^0, only material which can stand in I^0 can move. Therefore, main verbs in English do not undergo this process, and instead the nearly meaningless auxiliary verb *do* is used.

(33) a. * Eat you pickles?
 * Which pickles eat you?
b. Do you eat pickles?
 Which pickles do you eat?

Where does I^0 move to? The common answer to this question, based on work of the Dutch linguists Jan Koster and Hans den Besten,[1] is that it is C^0. The reason for this speculation is that the rule involved here seems to be in complementary distribution with complementizers. That is, its effects are blocked in those situations where we believe that complementizers are present in this position. Thus, there is a distinction between (34a) and (34b).

(34) a. Have you eaten?
b. * I remember (that) have you eaten.
c. Which pickles have you eaten?
d. * I remember that which pickles have you eaten.

Instead, Yes/No questions that are embedded are marked with a special complementizer: *whether*. And wh-questions in embedded contexts involve no evidence of a C^0 at all.

Under this view, the rule involved in these questions, then, might be formulated as in (35).

(35) I-to-C Movement
Adjoin I^0 to C^0.

The Word Criterion will correctly block movement of I^0 to C^0 in those cases where C^0 is filled with a complementizer, assuming that a complementizer and the modal or inflected verb in I^0 cannot together form one word.

Okay, so this is, roughly, a sketch of the rules that go into making up our knowledge of this fragment of English syntax. We've got three rules, one that

1 See den Besten (1983) and Koster (1975).

moves auxiliary verbs to I^0, another than moves I^0 to C^0 and a third that inflects main verbs. The first and last are driven by the Stray Affix Filter, or some parallel condition on the view that inflection is assigned under AGREE, and the second arises in questions. Both are subject to the Word Criterion, which prevents I^0, C^0 or V^0 from combining whenever they cannot form a word.

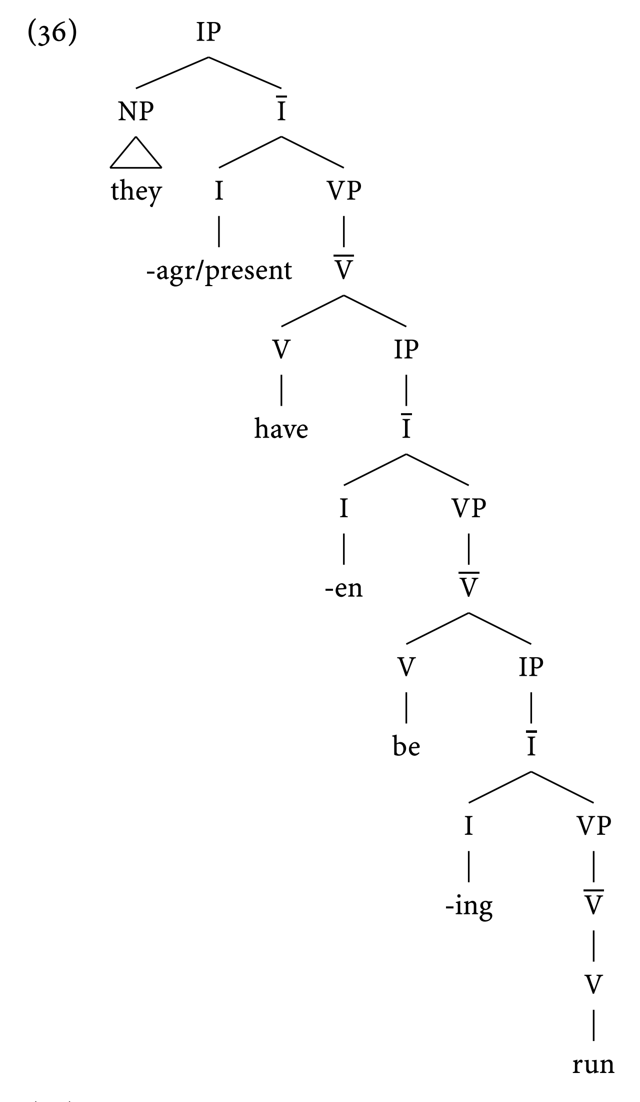
There's one last feature of these rules that we should consider. In all the cases so far examined, the V^0 that moves to I^0 is always the one that is closest to it. It appears that this is not just an accident of the examples we have chosen, instead it looks like it is a necessary feature of the verb movement rule. Consider, for example, how these rules might combine to apply to a representation like (36) below. Suppose that we move *be* to *-en* and form thereby the participle form: *been*. Imagine, further, that the I^0 which determines the present participle, the I^0 holding *-ing*, inflects the main verb that follows in whatever way it turns out that main verbs inflect. All that would then remain is to provide the tns/agree morphology associated with the highest I^0 with a stem. The correct outcome is one in which *have* moves to this I^0. But what would prevent the participle *been*, residing in the I^0 directly beneath *have*, from moving into this I^0. The outcome would be (36), and this is clearly not possible.

(37) * The been-re have running.

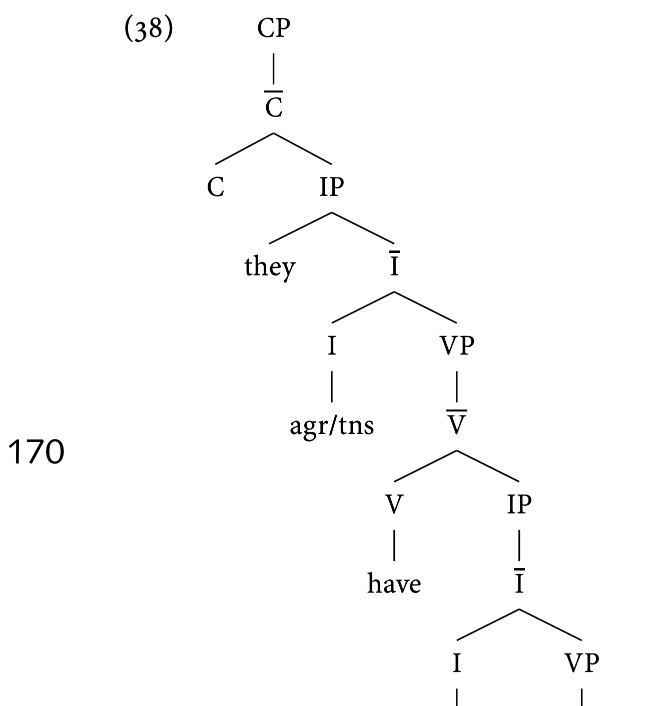
There's not even an inflectional form for *be* in which agreement/tense inflection combines with the (perfect) participle.

Similarly, consider how our rules might combine to apply to the D-structure representation like (38). In this example, we might imagine that, as before, the I^0 associated with *-ing* inflects the following main verb and that *be* moves into the I^0 associated with *-en* and forms the participle: *been*. Let verb movement also bring *have* into the I^0 associated with tense/agreement morphology to form the finite verb *have*. Finally, assume that this CP is a question and so triggers the rule that brings I^0 into C^0. We might imagine that the I^0 holding *been* could move into C^0; but this is not grammatical, as (39) indicates.

(36)


IP
NP
Ī
they
I
VP
-agr/present
V̄
V
IP
have
Ī
I
VP
-en
V̄
V
IP
be
Ī
I
VP
-ing
V̄
V
run

(38)


CP
C̄
C
IP
they
Ī
I
VP
agr/tns
V̄
V
IP
have
Ī
I
VP

(39) * Been they have running?

There might well be independent reasons for some of these effects. It might be, for instance, that English morphology does not allow participles to inflect for tense and agreement. And yet, all of the blocked outcomes fit a generalization that it might be worthwhile crediting for the whole class of cases. Travis (1984) has made just such a proposal. She suggests that there is a constraint on movement rules that relocate X^0s that blocks them from moving past other X^0s. She calls this condition the "head movement constraint," and I will formulate it as (40).[2]

(40) The Head Movement Constraint
No X^0 may move past a Y^0 that c-commands it.

As we will see, this feature of the grammar of English verb placement can be assigned to Universal Grammar. The Head Movement Constraint seems to govern instances of X^0 Movement in other languages.

4.3 Verb Second word order

Let's now consider verb placement phenomena in languages closely related to English. We'll start with examples from the Germanic family. In German we find that the position of the finite verb depends, roughly speaking, on whether the clause it is in is embedded or not. In embedded clauses, the verbs stack up at the end of the sentence in the mirror image of their arrangement in English: the finite verb comes last.

(41) a. . . . daß Hans das Buch kauft.
. . . that John the book buys

b. . . . daß Hans das Buch gekauft hat.
. . . that John the book bought has

c. . . . daß Hans das Buch gekauft haben muß.
. . . that John the book bought have must.

From this we can conclude, perhaps, that VPs are head-final. Further, if there are IPs and a method of combining verbs with the inflection associated with

2 She has a different formulation which builds in more context sensitivity and relates it to the notion of government that was described in chapter 3. For the cases that we will look at, (40) will do.

the I^0 of these IPs that employs Head Movement, they too are head final in German.

Interestingly, we find a different word-order in root, or independent, clauses. Here, the inflected verb no longer comes finally in the series. Instead it appears immediately following the subject. So unlike (41), we find only the word order in (42).

(42) a. Hans kauft das Buch.
John buys the book

b. Hans hat das Buch gekauft.
John has the book bought

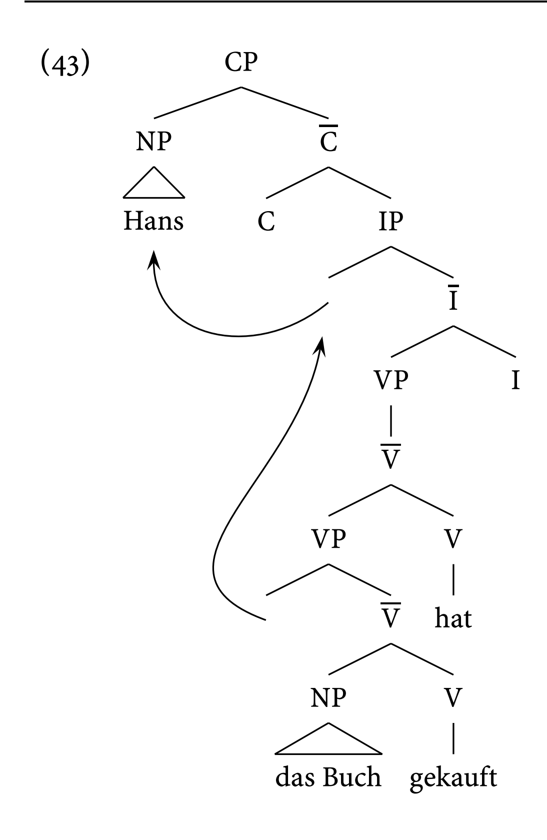
c. Hans muß das Buch gekauft haben
John must the book bought have

This seems to suggest that there is a movement rule which relocates finite verbs into the position immediately following the subject. Using the logic of the correlation argument, we might imagine that the position where finite inflection is in German immediately follows the subject, and it's into this position that finite verbs are driven in German.

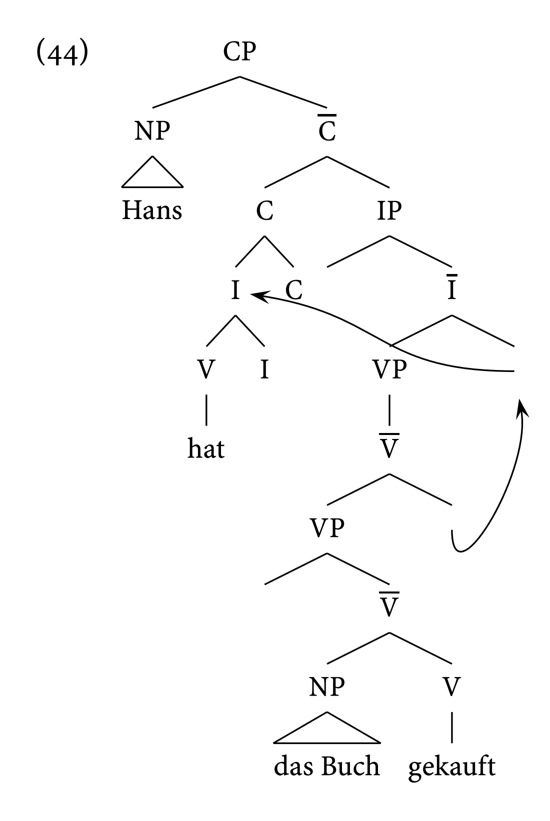
But this would miss the fact that the position of the finite verbs differs for embedded and independent clauses. What we want is some way of forcing verbs to move into the post-subject position in root clauses only. This suggests that it is not the finite distinction that is responsible for verbs' position in root clauses, but something else. Something that distinguishes root from embedded clauses.

We've already seen a similar difference in the grammar of English: recall that I^0-to-C^0 movement is restricted to root clauses in English. Perhaps the verbs are moving through I^0 into C^0 in cases like (42), then. This would credit German and English with two differences. On the one hand there is the difference in headedness that we see most directly in embedded contexts. And then there is something that allows/forces subjects to move past C^0 in embedded clauses. We might imagine that this second force, whatever it is, is like the process that moves wh-phrases in English into Specifier of C^0. Thus, the examples in (42) might get a parse like that in (43) on the next page. The subject moves into Specifier of IP, to satisfy the EPP, and then into Specifier of CP. The finite verb follows the subject because it moves through I^0 into C^0, as indicated in (44) below.

(43)

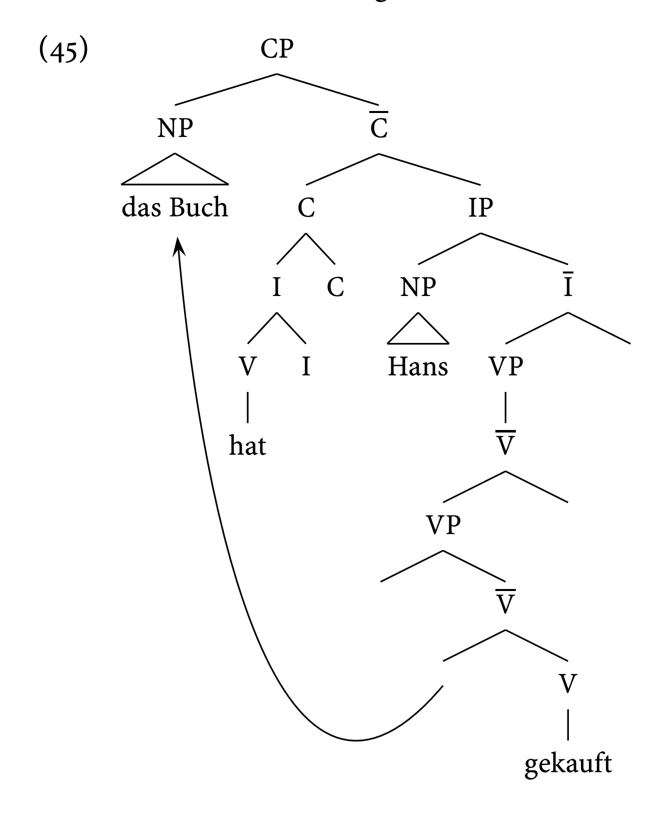

CP
NP
C̄
Hans
C
IP
Ī
VP
I
V̄
VP
V
V̄
hat
NP
V
das Buch
gekauft

(44)


CP
NP
C̄
Hans
C
IP
I
C
Ī
V
I
VP
hat
V̄
VP
V̄
NP
V
das Buch
gekauft

This leads to the expectation that we might find other phrases preceding the finite verb; and we do. It is not just the subject that may immediately precede the finite verb in root contexts, any phrase can. When some other phrase comes before the verb, the subject (typically) immediately follows the finite verb and the phrase that shows up at the front is understood to be "topicalized." We can capture these word-orders by letting the finite verb move into C^0 in all these clauses, but letting the Specifier of CP be filled by any phrase through a movement rule. Thus, alongside (43) we also find (45).

(45)



The generalization about German word-order can be described this way: any phrase may be initial, but exactly one must be. German is sometimes described as having "Verb Second" word-order for this reason. The account we've just sketched of this captures Verb Second by way of a rule that moves the verbs that have moved into finite I^0 into C^0, and by moving anything, but something, into the single position that exists to the left of C^0.

That the verb in these cases has in fact moved into C^0 is further substantiated by cases where Verb Second word-order is found in embedded clauses. Though there is considerable dialectal variation here, in the standard dialects,

V2 is possible in embedded clauses just in those cases where a complementizer may go missing. As in English, it is possible in German to unexpress the complementizer when the CP it heads is the complement to a certain class of verbs. The verb *say*, for example, can go without a complementizer — as in (46) — and when it does so in German we find V2 word-order, see (47).

(46) Jerry said (that) Sally has a dime.

(47) a. Er sagt, daß die Kinder diesen Film gesehen haben.
He says that the kids this film seen have
'He says that the kids have seen the film.'

b. Er sagt, diesen Film haben die Kinder gesehen.
he says this film have the kids seen
'He says that, this film, have the kids seen.'

Despite the similarities in verb movement that German and English have, note that one place where they differ is whether main verbs fall under the scope of verb movement. Unlike English, German main verbs can undergo movement. This is seen by their ability to move into C^0 in main clauses, as in (42a).

There are a host of mysteries about the grammar of verb placement in German that we will not attempt to solve here. What, for instance, is responsible for forcing movement of something into the Specifier of CP in German root clauses. And why isn't the movement of a phrase into Specifier of CP allowed in embedded clauses, as in (48)?

(48) * Ich sagte [$_{CP}$ das Buch [daß [$_{IP}$ Hans [$_{VP}$ gekauft] hat]]].

Vikner (1995) discusses some of the ideas that have been offered for answers to these questions.

There are some things that we might notice about the verb movement processes in German which are similar to the parallel processes in English. For example, the Head Movement Constraint is preserved in the grammar of verb placement in German. Sentences like the following, where *haben* ('have') moves past the modal are ungrammatical in German just as they are in English.

(49) * Hans haben das Buch gekauft muß
John have the book bought must
.

Moreover, we can observe that there is a pattern to all these rules in German and English: the moved term always adjoins to another head. Because this appears to be something that is the same across German and English, it is a candidate language universal. Indeed, over a large range of cases examined across

languages, this generalization appears to hold up. This suggests a constraint movement rules that is sensitive to the phrase/head distinction. One of the first proposals for such a constraint is found in Baltin (1982), who argues that this is part of a more general constraint on movement rules. He proposes that the position to which a phrase or head is moved always has the same phrasal status as the phrase or head being moved. So a head adjoins to another head position only, and a maximal projection can only move to another maximal projection position, and so on. Let's call this the Like-attracts-Likes Condition:

(50) Likes Attracts Likes
An X^0 may only adjoin to, or substitute into, a position that is also an X^0, and an XP may only adjoin to, or substitute into, a position that is an XP.

Finally, note that with the sole exception of Affix Hopping, the movement rules we have examined in these two languages all have the feature that the verb or I^0 that has moved has moved up. We don't find cases where Verb Movement has relocated the verb downwards, as in examples like (51).

(51) a. * ...daß Hans das Buch gehabt kauft.
...that Hans the book had buys

b. * John must had buy the book.

We have already found that this is a feature of the Argument Movement rule — it also relocates terms only to a position higher in the phrase marker.

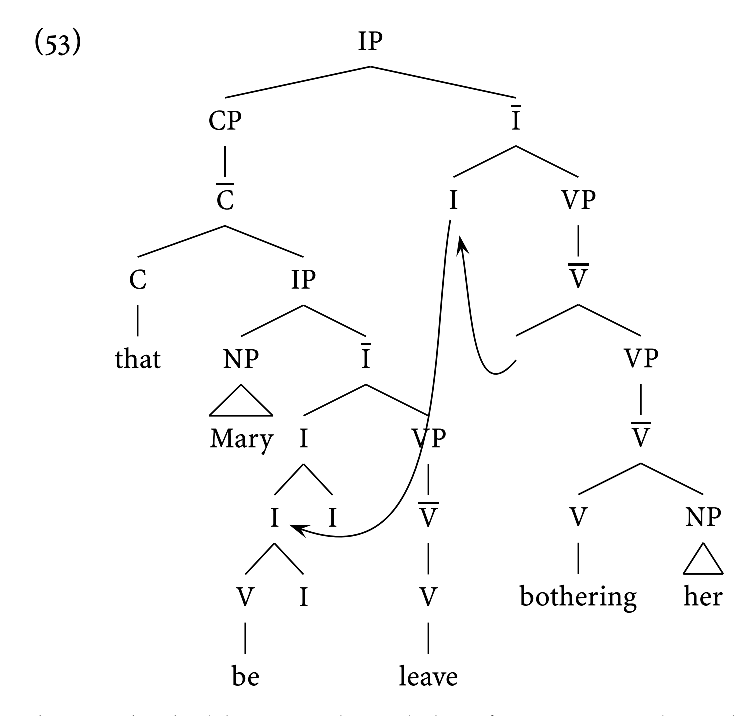
Actually, we found that the constraint was more specific than this; it required that the moved term relocated to a c-commanding position. Let us formulate this constraint, shared by both German and English Verb Movement rules, as follows.

(52) Upwards Constraint
α can move to position β only if β c-commands α.

That this narrower condition is required for verb movement is shown by cases like (53), which would be possible if verbs could move to non-c-commanding positions.[3]

3 This parse does not reflect the topicalization that we've seen evidence for — recall, there is reason to believe that finite "subject" clause has been topicalized to IP.

(53)



This too looks like a good candidate for a universal condition: things only move to c-commanding positions.

Let's take a look now at what we find with respect to verb placement in some of the other Germanic languages. In the Scandinavian languages we find a situation similar to German, as Vikner (1995) reviews. I will use Danish as a guide; much of what we see for Danish is found in Norwegian and Swedish as well.The same sensitivity to embedding is found in the placement of the finite verb in these languages. As the contrast in (54) indicates, Danish is like German in allowing any constituent, but only one, to precede the finite verb in independent clauses. That is, it shares with German the trait of being "Verb Second."

(54) a. Børnen har set denne film
kids-the have seen this film
'The kids have seen this film.'

b. Denne film bar børnen set.
this film have kids-the seen
'The kids have seen this film.'

c. * Denne film børnen har set.
this film kids-the have seen
'The kids have seen this film.'

But, as (55) shows, these traits are not shared by dependent clauses, where, instead, the subject must precede the finite verb.

(55) a. * Jeg ved ikke hvor i går har koen stået.
I know not where yesterday has cow-the stood
'I don't know where the cow stood yesterday.'

b. Jeg ved ikke hvor koen har i går stået.
I know not where cow-the has yesterday stood
'I don't know where the cow stood yesterday.'

Moreover, as the contrasts in (56) show, the placement of the finite verb relative to negation is sensitive to the embedded/non-embedded context.

(56) a. Børnen har ikke set denne film.
kids-the have not seen this film
'The kids haven't seen this movie.'

b. * Børnen ikke har set denne film
kids-the not have seen this film
'The kids haven't seen this movie.'

c. Jeg ved at børnen ikke har set denne film
I know that kids-the not have seen this film
'I know the kids haven't seen this movie.'

d. * Jeg ved at børnen har ikke set denne film
I know that kids-the have not seen this film
'I know the kids haven't seen this movie.'

This indicates that the finite verb has moved out of VP, past negation, into C^0 in independent clauses, just as it does in German.

It looks, therefore, like Danish has the word-order of English — placing verbs before their objects — but the syntax of verb movement that German has. Note in particular that main verbs are able to move as well as auxiliary verbs, as we see partly in (55) (look at the root clause word-order). If we were to look further, we would see that Danish (and the remainder of Germanic) also have the range of constraints we have seen on verb movement operations.

There are a couple differences, however. One at least potential difference between Danish and German is the existence of V^0-to-I^0. Whereas we cannot easily discern whether such an operation exists in German, it apparently does not in Danish (nor does it in standard Norwegian and Swedish), since when the verb has not relocated into C^0, it remains to the right of negation. (Of course,

we have made the perhaps incorrect assumption that negation occupies the same place in Danish as it does in English.)

Another difference concerns the range of contexts where verb second is found in embedded contexts. Recall that in German the verb moves into C^0 in embedded clauses only in those contexts where a complementizer is not required to fill that C^0. But in Danish (as in the other Scandinavian languages), V2 word-order is possible even in those embedded clauses which have a complementizer associated with them, as in (57).

(57) Vi ved at denne bog har Bo ikke læst.
we know that this book has Bo not read
'We know that Bo hasn't read this book.

A variety of hypotheses about this situation have been offered — the chapter following — the assigned one in Vikner gives a good overview; let us adopt for concreteness the hypothesis that there is a CP "shell" that can be embedded within a regular CP. This CP shell provides the C^0 into which verbs move in embedded clauses in Scandinavian.

So, let's see what we've got so far. If we factor out the parts to the set of rules in German, Dutch and English that are common to the syntax of verb placement, we have the following:

(58) Universals
a. Likes Attracts Likes
b. Upwards Constraint
c. The Head Movement Constraint
e. Word Criterion

The differences that are left can be boiled down to the following four parts.

(59) a. German VPs (and IPs?) are head final.
b. English and Danish VPs (and IPs?) are head initial.
c. German and Danish: verbs (and I^0?) are movable.
d. English: auxiliary verbs and I^0 are movable.
e. German and Danish: Root C^0 and Specifier of CP must be filled. (i.e., V2)
f. English: Only wh-phrases move to Specifier of CP, and only questions trigger I^0-to-C^0 movement.

Putting this in the terms of language acquisition which we started with, the differences in (59) constitute what the child must learn in order to determine whether the grammar he or she is acquiring is German, Danish or English.

4.4 The Pollockian revolution

Let us now add to our family of languages: French. In French we also find evidence of rules that move verbs. Just as in English, French verbs distribute themselves relative to polarity items, like sentential negation, on the basis of their inflectional class. Thus, for example, we find that finite verbs in French, just like their counterparts in English, only appear before negation.

(60) a. Jean n'a pas lu livres.
John n'have not read books
'John hasn't read books.'

b. * Jean ne pas a lu livres.

A word about negation in French is needed before the contrast in (60) will be interpretable. Unlike English, sentential negation in French is expressed by way of two words, one (*ne*) appears to be in a position like that we have put inflectional morphemes in English. We might imagine that like the contracted form of *not* in English (*n't*) it is contracted onto whatever word has moved into I^0 (as in (60a)). The second word (*pas*) is the one that has a distribution more like that of English not. And, as (60) illustrates, the finite form of have (*ha*) must precede this part of sentential negation.

French differs from English (but is like German) in moving main verbs. This can be seen in French by the fact that they are placed to the left of sentential negation when they are finite, as (61) demonstrates.

(61) a. Jean n'aime pas Marie.
John ne'love not Mary
'John doesn't love Mary.'

b. * Jean ne pas aime Marie.

As the contrast between the position of main verbs in (61a) and (60a) indicates, it is the inflectional class of the verb that determines its position. That is, just as in English, there is a correlation between the inflectional class of the verb and its syntactic position — a correlation that is captured by fixing the position of inflectional morphemes with the phrase structure rules and driving the verbs to these positions with the verb movement operation.

That main verbs may move is also responsible for another difference between French and English. This derives from the ordering that the Projection Principle places on arguments and non-arguments. As we have had occasion to discuss several times, now, the Projection Principle forces arguments to be closer to their θ-marking head than it will allow non-arguments. In verb initial languages, like French and English, this means that complements should come before non-complements. This is roughly true for English, as we have seen; but it is not the case for French.

(62) a. Jean embrasse souvent Marie.
b. * John kisses often Mary.

This apparent difference in the function of the Projection Principle can actually be seen as a product of verb movement. Since we already know that main verbs in French but not English move into I^0, it makes sense that the main verbs in French, but not English, should be able to be separated from their objects by all sorts of material, including adverbs.

But now consider what happens in French non-finite clauses.[4]

(63) a. **Comprendre** *à peine* l'italien apres cinq ans d'étude
understand barely the-Italian after five years of study
dénote un manque de don pour les langues.
shows a lack of gift for the languages
'To barely understand Italian after five years of study shows a lack of talent for languages.'

b. **Perdre** *complètement* la tête pour les belles étudiantes c'est
lose completely the head for the pretty students it is
dangereux.
dangerous
'To completely lose your head for pretty students is dangerous.'

(64) a. ne *pas* **sembler** heureux est une condition pour écrire
ne not seem happy is a condition for writing
des romans
novels
'To not seem happy is a (pre?)condition for writing novels.'

4 These data all from Pollock (1989).

b. * ne **sembler** *pas* heureux est une condition pour écrire
ne seem not happy is a condition for writing
des romans.
novels
'To not seem happy is a (pre?)condition for writing.'

Here it appears that main verbs may move past adverbs but not negation. (The infinitival verbs are in bold face, in these examples, and the term that separates them from their objects is in italics.) Auxiliary verbs behave differently.

(65) a. ne *pas* **être** heureux est une condition pour écrire des romans
ne not be happy is a condition for writing novels
'To not be happy is a condition for writing novels.'

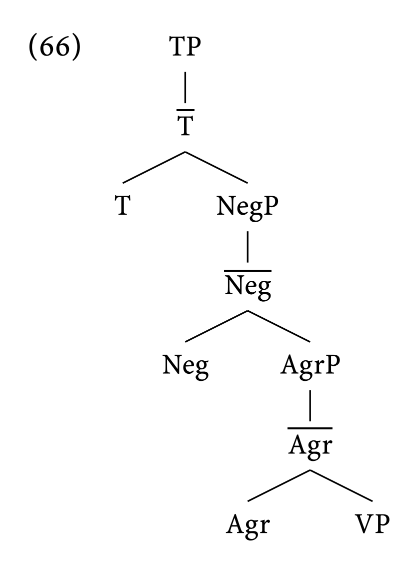
b. N'**être** *pas* heureux est une condition pour écrire des romans.
ne'be not happy is a condition for writing novels
'To not be happy is a condition for writing novels.'

They may optionally move past negation.

We learn two important things from these facts. First, that the contrast between auxiliary and main verbs that seems to distinguish English from German/Danish is actually not a distinction in languages. Instead, it is a difference which is found within a single language: French. That is, this is not a parameter along which languages vary, then, but rather a parameter along which clause-types vary. We need to express the distinction between English, and these other languages, in terms of the clause types these languages host, not in terms of the targets for their verb movement rules. So, we're looking for something that distinguishes English finite clauses and French non-finite clauses on the one hand from French finite clauses on the other.

The second lesson of this paradigm is the main point of Pollock's paper. He suggests that the two different positions that main verbs may occupy across finite and non-finite clauses warrants giving a structure like that in (66) to clauses, where the higher I^0 is equated with Tense and the lower one with Agreement. This answers to the correlation that appears to hold for main verbs between whether they bear tense morphology or not: in finite clauses they do, and in non-finite clauses they do not.

(66)



Pollock also introduces the idea that Neg heads a phrase which stands between the two other Inflectional phrases, as indicated in this pares. There are several reasons for wanting to treat negation differently from other adverbials. One is that the syntactic position of negation is comparatively rigid when compared to other adverbs. Another is that only negation blocks the process that inflects main verbs in English, as we saw earlier; other adverbs don't. Pollock suggests distinguishing negation from adverbs structurally, and then making reference to this structural difference to capture these ways in which negation behaves uniquely. I won't examine in detail how this idea works, partly because we will eventually go in a slightly different direction than does Pollock. But let us nonetheless adopt — at least as an intermediary hypothesis — the thesis that negation does head a phrase as shown in (66). It should be noted, however, that this introduces a problem: how is it that Agr^0 can move to T^0 past negation without violating the Head Movement Constraint. Pollock offers a solution to this problem that I will come to soon.

Note that though the difference between main and auxiliary verbs in French non-finite clauses that we've just reviewed speaks on behalf of two head positions to which verbs may move, it doesn't really indicate what the identity of these two head positions might be. While most of Pollock's proposal involves examining how the hypothesis that there are two inflectional positions within a sentence can be used to explain the differences in verb position across clause types, he also assigns values to these inflectional phrases. But, in fact, it is extremely difficult to tell what the value of these heads is. Pollock decides in favor of giving the higher I^0 the value of Tense and the lower one the value of Agree-

ment. This involves two decisions: first that these positions should be equated with inflectional categories, and second that they should be equated with these particular ones. The first decision, note, is a bit of a leap here as there is no convergent evidence of the headedness sort to anchor this conclusion. That is, part of the reason we decided that the landing site of verb is I^0 came in response to the evidence that I^0 was a head.

There is, however, a sort of reasoning which suggests that the first decision is correct, but that the second isn't. Consider the following verbs, drawn from a variety of IndoEuropean languages.[5]

(67) a. Legg - eva - no
read - Imperfect - 3plur

b. Parl - er - ò
speak - Future - 1sing

c. Buhl - t - en
woo - past - plur

Note that in each of these cases, the morpheme which encodes tense information precedes that which encodes agreement. Is there a way of determine the syntactic arrangement of inflectional terms from their morphological arrangement? Mark Baker has made famous an argument for thinking that the answer to this question is yes. Following proposals by Pieter Muysken and Donna Gerdts, Baker argues that there are correlations between the order that inflectional affixes have relative to the verbal stem and the order of the syntactic operations that these affixes encode. Thus, for example, in Chamorro, the passive morpheme necessarily comes closer to the verbal stem than does the subject agreement affix, as (68) shows.[6]

(68) Para.u.fan - s - in - aolak l famgu'un gi as tat-n-niha
irr.3pS - pl - pass - spank the children obl father.their
'The children are going to be spanked by their father.'

This ordering on affixation mimics the order in which the syntactic operations of Passive and subject agreement take place — the passive must bring the underlying object into the subject relation before the agreement process can apply.

5 The argument I'm reporting here comes from Belletti (1990).

6 These data come from Gibson (1980). The principle formulated in Baker (1985) has its roots in many earlier works, among which are: Muysken (1979, 1981) and Gerdts (1981).

Baker calls correlations of this type, which he suggests are widespread enough to be considered law-like, the "Mirror Principle."

(69) MIRROR PRINCIPLE
Morphological derivations must directly reflect syntactic derivations (and vice versa).
(Baker 1985, p. 375)

Baker argues that the Mirror Principle will emerge once the correct theory of the syntax-morphology interface is found. In particular, (69) can be derived on models that posit for each of the agreement, passive, and other such operations one rule underlying both the morphological and syntactic effects. Baker tentatively concludes that these metarules apply in the Syntax — that is, after the verbal roots are inserted into phrase-markers. The resulting picture is sketched in (70).

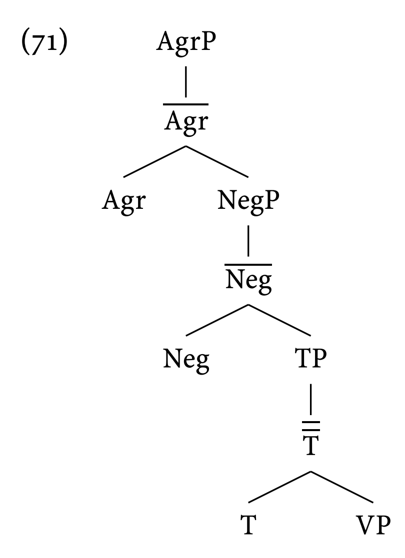
(70)

$$\begin{array}{lcl} \text{D-structure:} & V_{root} & \\ & \downarrow & \leftarrow Metarule\,\alpha \\ & V_{root} + Af_{\alpha} & \\ & \downarrow & \leftarrow Metarule\,\beta \\ \text{S-structure:} & V_{root} + Af_{\alpha} + Af_{\beta} & \end{array}$$

Each metarule brings about the relevant syntactic operation and adds to the verbal root the corresponding affix. Note in particular that this picture correlates syntactic operations with affixes; the correspondences that Baker summarizes with the Mirror Principle only concern the relation between affixal orderings and the orderings of syntactic operations. Indeed, Baker's argument leads to the conclusion that the metarules in (70) necessarily involve affix-syntactic operation pairs. It is the order of affixes that correlates with the relevant syntactic operations, and not some more abstract information, such as morphological features or the like. We shall have an opportunity to revisit this point.

Now if Baker's conclusions from the Mirror Principle are imported into the domain of Verb Movement, then the arrangement of morphemes in (67) suggests that the higher of Pollock's positions should be associated with agreement morphology, and the lower with tense morphology. Then, the fact that tense morphology comes closer to the verbal stem will follow from the Head Movement Constraint. So we should adopt, perhaps, a picture like that in (71) below.

(71)



We can reach this conclusion, perhaps, through a different route. If we compare Icelandic with Danish, we see a difference in verb placement that suggests that only Icelandic has movement of verbs into the I^0 that stands above negation. In embedded Icelandic clauses the finite verb must precede negation, as in (72).

(72) a. * Ég spurði af hverju Helgi ekki hefði lesið þessa bók
I asked whether Helgi not had read this book
'I asked whether Helgi hadn't read this book.'

b. Ég spurði af hverju Helgi hefði ekki lesið þessa bók
I asked whether Helgi had not read this book
'I asked whether Helgi hadn't read the book.'

But as we have seen, Danish finite verbs cannot precede negation. Another example illustrating this fact is (73).

(73) a. Jeg spurgte hvorfor Peter ikke havde læst den.
I asked why Peter not had read it
'I asked why Peter hadn't read it.'

b. * Jeg spurgte hvorfor Peter havde ikke læst den.
I asked why Peter had not read it
'I asked why Peter hadn't read it.'

If we assume that the position of negation is constant across these languages, then this indicates that the verb has moved past this position in Icelandic, but not Danish. Now, interestingly, Icelandic has a full paradigm of subject agreement, but Danish has no subject agreement. Comparing Icelandic with Danish,

then, what we find is a correlation between subject agreement and position to the left of Neg^0. Indeed, this correlation holds throughout the Scandinavian languages.[7] Standard Swedish and Norwegian are like Danish in lacking subject agreement morphology, and they are like Danish as well in placing their finite verbs to the right of negation. This is just the correlation that (71) predicts.

Adopting this modified Pollockian parse for clauses, we are left with a series of questions about how verbs distribute themselves over the two head positions. Why can main verbs not move to the "agreement" X^0 in non-finite clauses in French, for instance; and why are they unable to occupy this position even in finite clauses in English? I don't think we have entirely satisfactory answers to these questions. Pollock relies on an observation due originally to Ian Roberts: The X^0s in which main verbs cannot surface are devoid or close to devoid of morphological content. English agreement is much less "robust," somehow, than is French agreement. And subject agreement in French non-finite clauses is considerably less pronounced than it is in French finite clauses. (In fact, it is thoroughly absent.) Now, this difference in agreement morphology Pollock exploits to control whether or not a verb can move into the lower of the two inflectional X^0s – this is because Pollock assigned agreement to this lower head. But we can copy his idea into the assignment we have given to these positions. Tense morphology is also richer in French than it is in English, in both finite and non-finite clauses; and agreement is richer in finite clauses than it is in non-finite clauses too. In fact, there is no evidence of agreement at all in infinitival clauses.

Even if this "robustness" in the morphology associated with the T^0 and Agr^0 positions is not what is relevant (and further cases suggests that it isn't), let's continue to exploit the language that Pollock adopts for whatever the relevant difference turns out to be. Those heads that are associated with morphology that allows main verbs to occupy them, let's call "strong." And those head positions that are not associated with morphology that allows main verbs to sit in them, let's call "weak." We can now describe in these terms what we've seen so far. In particular, we will ascribe to the Tense and Agreement positions

7 More or less – there are some cases in which agreement morphology exists, though in a very impoverished form, and verb movement past negation nonetheless appears to be blocked. (This happens in certain dialects of Faroese, in a register of Norwegian, and in a Swedish dialect heavily influenced by Finnish.) Some of these facts are reported in the Vikner (1995), where references to the relevant literature can also be found.

different "strengths," and in this way describe how they control the Move X^0 operation. For English and French, we might assign the values as shown in (74) and (75) below.

(74) **English:**

a. Finite:
Tense: weak
Agreement: weak

b. Infinitive:
Tense: *to*
Agreement: empty?

(75) **French:**

a. Finite:
Tense: can't tell
Agreement: strong

b. Infinitive:
Tense: strong
Agreement: empty

In these value assignments, I've decided to treat the infinitival marker in English, i.e., *to*, as something that can optionally move from T^0 to Agr^0. On this view it is rather like an auxiliary verb in French, and should behave similarly, therefore, in infinitival clauses. For some speakers of English there is, in fact, an alternation that resembles that we've seen for French auxiliary verbs:

(76) a. Jerry tried [not to eat candy].
b. Jerry tried [to not eat candy].

Note that we cannot determine whether finite T^0 is strong or weak in French because verbs are always driven up to Agr^0 in this context.

Turning now to Scandinavian, we have something like:

(77) **Danish, Norwegian and Swedish:**

a. Finite:
Tense: weak
Agreement: not there

b. Infinitive:
haven't looked yet (it varies across the group, as it turns out)

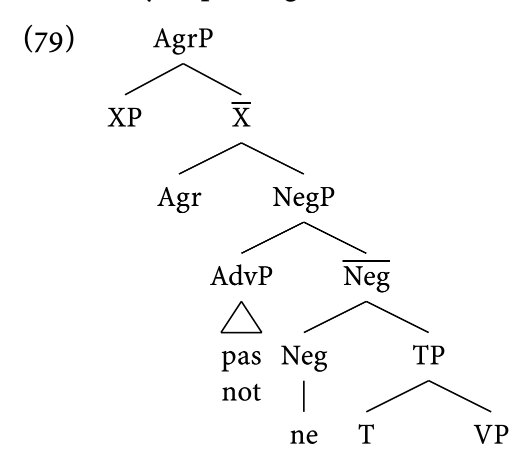
(78) **Icelandic:**

a. Finite:
Tense: can't tell
Agreement: strong

b. Infinitive:
haven't looked yet

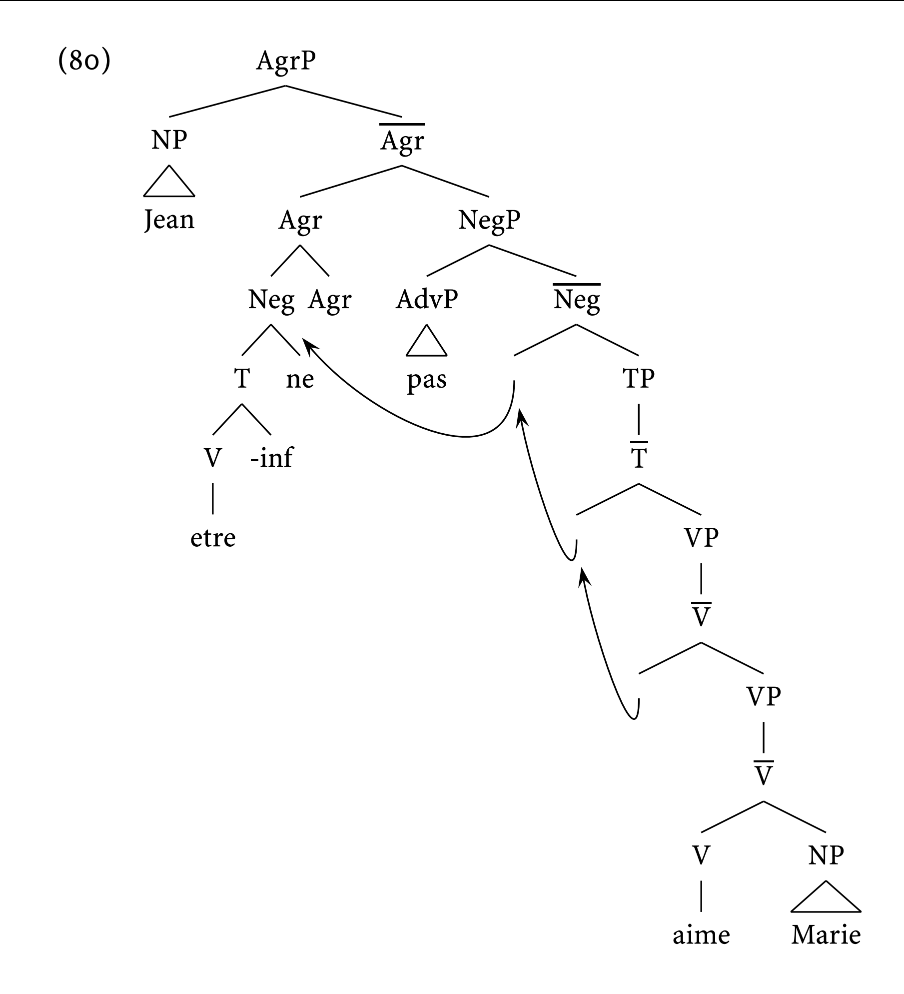
My decision to classify the Mainland group as having weak tense morphology is based on the fact that this language group behaves like English, and not like French, with respect to the positioning of main verbs relative to their complements. That is, in this group, non-complements cannot intervene between main verb and NP objects (except in contexts where the verb has relocated into C^0).

Okay, this is one part of the story. The morals are: (i) there are two inflectional heads between C^0 and VP in some languages, and (ii) it is properties of those heads that determine what must and can move into them. Let's now address the problem that treating negation as a head poses. Recall that we should not expect the Head Movement Constraint to allow movement of T^0 to Agr^0 past negation, if negation is a head that lies between them. Pollock's suggestion is that *not* and *pas* are in the Specifier of NegP, and that *ne* is what heads NegP in French. In English, we might imagine that this head is silent. Thus, the correct way of parsing sentences with sentential negation in them is as in (79).

(79)


On this view, then, when T^0 moves to Agr^0 it doesn't have to skip Neg^0, it can move through this position and pick up the head (if there is one) on the way. In a finite French clause, then, the surface parse might look roughly as in (80) on the next page. As this parse indicates, *ne* gets carried to the position before

(80)



pas by the tensed verb as it moves into Agr^0.[8] Note that we have to let something force the bound morphemes, that is, the tense and agreement suffixes, to linearize to the right of the term that has adjoined to them, but force *ne* to linearize to the left of the term that has adjoined to it.

8 The S-structure in (80) is derived as follows (ignoring the movement of the subject). The main verb, *aime*, is moved into T^0, where it combines with the tense morpheme. The resulting tensed verb (that is, T^0) moves and adjoins to *ne*. the resulting negated verb (that is, Neg^0) moves into Agr^0, where the agreement morpheme combines with the verb. This surface parse, then, claims that negation — *ne* — is closer to the verbal stem than is the agreement morpheme, and this is somewhat surprising. We must also understand how to get the right linear order of *ne* and the inflected verb out of this representation.

This can't be the only way that *ne* manages to get to the left of *pas*, because *ne* shows up before *pas* in infinitival clauses as well, even when no verb has moved into Agr^0. In these situations, we might imagine that *ne* moves on its own into Agr^0. This, in fact, is what Pollock suggests. He hypothesizes that *ne* adjoins to Agr^0 on its own in much the same way that so-called "clitics" do. "Clitics" are prosodically weak pronouns that, in the Romance languages, surface in construction with a verb; more particularly, they form a prosodic word with a verb and are therefore adjacent to that verb (or adjacent to something else that is part of the verb). In finite clauses, it is the finite verb that clitics appear with, and in infinitival clauses, it is typically the infinitival verb that the clitic appears on. One way of capturing this distribution is let clitics adjoin to Agr^0, where they will come close enough to the finite or infinitival verb to be parsed with that verb prosodically.

Clitics, then, instantiate another instance of movement. Object clitics, for instance, satisfy the Theta Criterion and Projection Principle by way of occupying a position that is sister to the verb they are arguments of, but surface in Agr^0 position. This type of movement is called "Cliticization." Cliticization doesn't have the same properties that Head Movement has — it isn't subject to the Head Movement Constraint, for instance — and therefore we can treat it is as being independent of the rules we are investigating here.

If we adopt Pollock's suggestion, that *ne* cliticizes to Agr^0, we might be better poised to understand why it is linearized to the left of the verb+ tense+-agreement complex. Indeed, we could impose a general constraint on adjunction that would linearize the moved item to the left of the phrase it adjoins to. Something along the lines of (81), perhaps.

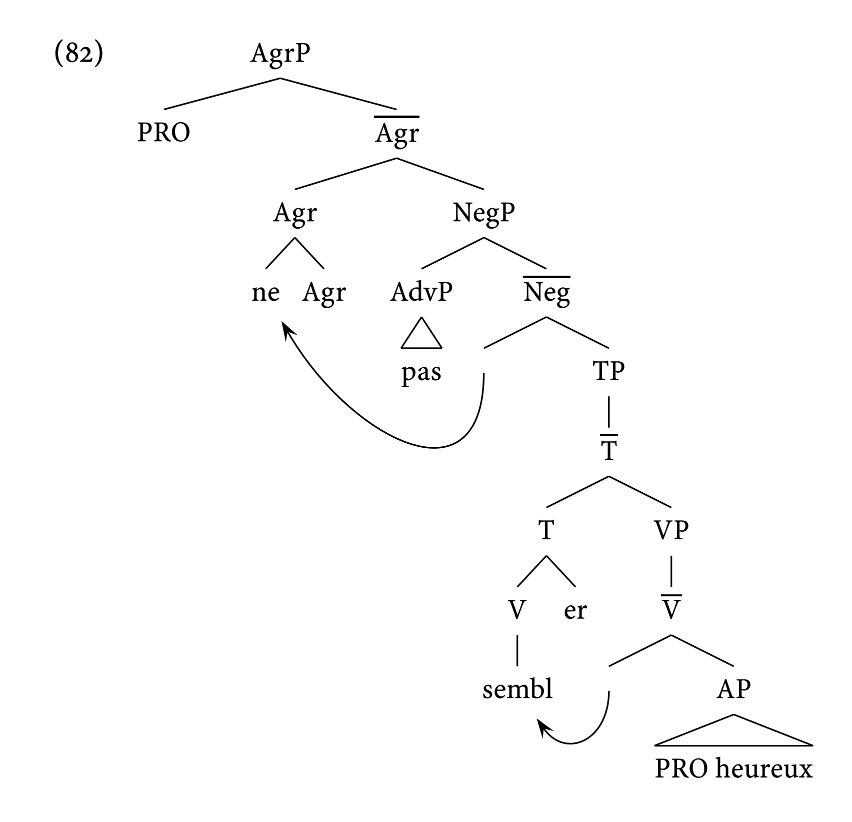
(81) Adjunction Linearization
If α adjoins to β, then α precedes β.

But it is not clear how this principle would achieve the right outcome from the derivation in which the verb collects *ne* on its way to Agr^0, as in (80). There is some unfinished business here, then. We will leave the matter of how material is linearized inside X^0s to a later chapter.

With this assumption about *ne*, then, an infinitive that doesn't have verb movement to Agr^0, but is negated, might have an S-structure like (82) on the facing page.

So, now what we would like to do is see if we can't get a better understanding of what it is about this "strong"/"weak" contrast that makes main verbs susceptible to it, but not auxiliary verbs. Pollock makes a proposal in this di-

(82)



rection. He suggests that we think of main verbs as differing from auxiliary verbs with respect to their ability to assign a θ-role. In particular, he suggests that we define main verbs as ones that have a θ-role to assign, and auxiliary verbs as ones that don't. He then crafts a proposal that would prevent verbs from assigning θ-role from a position associated with a "weak" inflectional head. And this will essentially prevent main verbs from both moving to and moving through the positions we have assigned the "weak" value. His idea goes something like this:

(83) a. Assume that movement operations leave a "trace" of the moved X^0 or XP in the position that the term is moved from.

b. Let the Theta Criterion be enforced at S-structure.

c. Let the trace of a verb be capable of assigning the verb's θ-role only if the verb has not adjoined to a "weak" X^0.

There are some immediate problems with this idea that Pollock himself addresses. It wouldn't seem able to account for the British (85), nor for the

contrast in (84).

(84) Mary hasn't a dime.

(85) a. Mary isn't unhappy.
b. * Mary seemsn't unhappy.

Why doesn't the *be* in (84a) have the same number of θ-roles as (84b)? And it sure looks like *has* in (85) is capable of moving past negation, though it appears equally like it assigns θ-role. Pollock suggests that the θ-role assigned in (85) comes not from the verb, but from some hidden term. Something similar might be attempted in distinguishing the cases in (84).

But there is another kind of problem, specifically for the prediction that movement through a weak X^0 is unavailable to main verbs, which seems to me insurmountable. Danish, as we have seen, and the other Mainland Scandinavian languages, show Verb Second word-order; (86) is an example, for instance, where the main verb has relocated to C^0 position.

(86) Købte han bogen
bought he book-the
'Did he buy the book?'

But we have also just discovered that Danish has weak tense morphology. Thus, if we preserve the Head Movement Constraint, this entails that Danish main verbs are capable of moving into T^0 and from there moving into C^0. If Danish T^0 is weak, this means, in turn, that moving through a weak X^0 is possible for main verbs.

For this reason, I will interpret the weak/strong typology to determine only whether a main verb may surface in the relevant X^0. This seems to be the way these values are interpreted most commonly in the literature. It is, in fact, on this basis that I gave the assignments of "weak" and "strong" to Agr^0 and T^0 in English, French and the Scandinavian languages in (74), (75), (77) and (78). In these assignments, then, I have given a "weak" value to those X^0 positions for which there is evidence that main verbs cannot surface, and to those X^0 positions where there is evidence that main verbs can surface I have given the "strong" value. I have also indicated certain positions — namely, French and English infinitival Agr^0 — as empty positions. This is my, somewhat idiosyncratic, way of expressing the optionality of movement to these spots. Because they have no morpheme in them, these infinitival Agr^0s do not cause the Word Criterion force a verb stem to move into this position. But, because it is present, it is still a possible landing site for X^0 movement. It's in this respect, then, that

the infinitival Agr^0s of these languages are different from the finite Agr^0 of, say, Danish. Movement of a verb to the position we've identified for finite Agr^0 is not possible in Danish — neither auxiliary nor main verbs may surface in this position — and I've coded that here by making Agr^0 completely absent in Danish.

Under this typology, X^0 positions identified with inflectional morphemes come in four strengths. In their weakest manifestation, they are completely absent, and nothing, as a consequence, may surface in their position. In their next strongest manifestation, they are present but are associated with no morpheme. In this guise we find optional movement of the infinitival marker *to* and auxiliary verbs, but main verbs are still not capable of surfacing in them. (This is what we see for the infinitival Agr^0 in French.) In their next strongest manifestation, they are associated with a "weak" morpheme. At this level of strength, the Word Criterion kicks in and requires that something express the morphology they are associated with. The "weak" property, however, prevents main verbs from doing this by adjoining to the morpheme. Finally, there are the X^0 positions occupied by a "strong" morpheme. These positions allow either auxiliary or main verb to surface in them, and the Word Criterion forces one or the other to. Because we want to prevent main verbs from surfacing in all but the last of these positions, the constraint that controls the auxiliary/main verb distinction should be stated as (87).

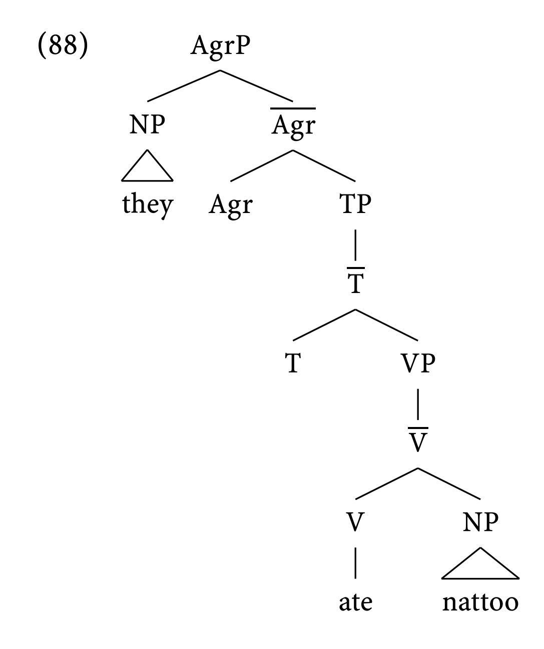
(87) A main verb adjoined to α at S-structure creates ungrammaticality unless there is a "strong" morpheme in α.[9]

4.5 AGREE

We've developed a rough description of how main and auxiliary verbs are driven into I^0s that uses the "strong"/"weak" distinction. We should now look at how this matter interacts with the way in which inflectional heads assign their morphology to the relevant verbs. The conception we have adopted of how verbs inflect is that it parallels the kind of relation that we have seen in connection with Case. The tense and agreement heads assign the relevant inflection to the verb in configurations like (88).

9 This is going to have certain consequences for the syntax of a sentence in which the verb has moved into C^0. In these cases — questions and, in the Germanic languages other than English, root finite clauses — we must imagine that there is a strong morpheme in C^0.

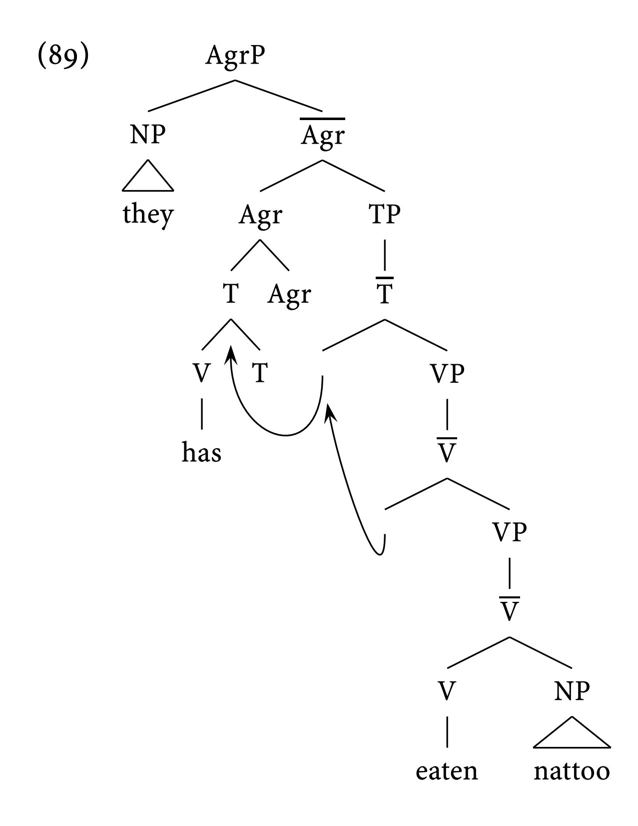
(88)



When auxiliary verbs are involved, the verb assigned inflection moves to the heads responsible for that inflection, as in (89) on the next page. The syntactic results of inflecting the verbs in these two cases is obviously different, and so a reasonable first question is: could the method of inflecting auxiliary and main verbs be different? Could it, for instance, be that main verbs get inflected by assignment while auxiliary verbs get inflected by moving?

This direction looks unlikely. The conditions we have placed on verb movement are mirrored by conditions on the assignment relationship that holds in the case of main verbs. Movement of V^0 and I^0 can move the term to just c-commanding positions, and never past a c-commanding X^0. And, similarly, the assignment relation always goes from an I^0 (or C^0) to a term that it c-commands, and it never crosses a c-commanding X^0. This is indicated by the fact that the parse in (90) cannot produce the outcome in (91). This would be possible if the Agr and T heads of the root clause were able to assign their inflection to the verb that heads the embedded VP. This would be prevented if these heads are prevented from assigning their inflection past the "closer" head: *make*. This is quite general: cases where an I^0 assigns the inflection it determines on a head that is embedded below another possible target for its inflection do not seem to arise. This looks completely parallel to the Head Movement Constraint, which proscribes movement of a head past a c-commanding head.

(89)



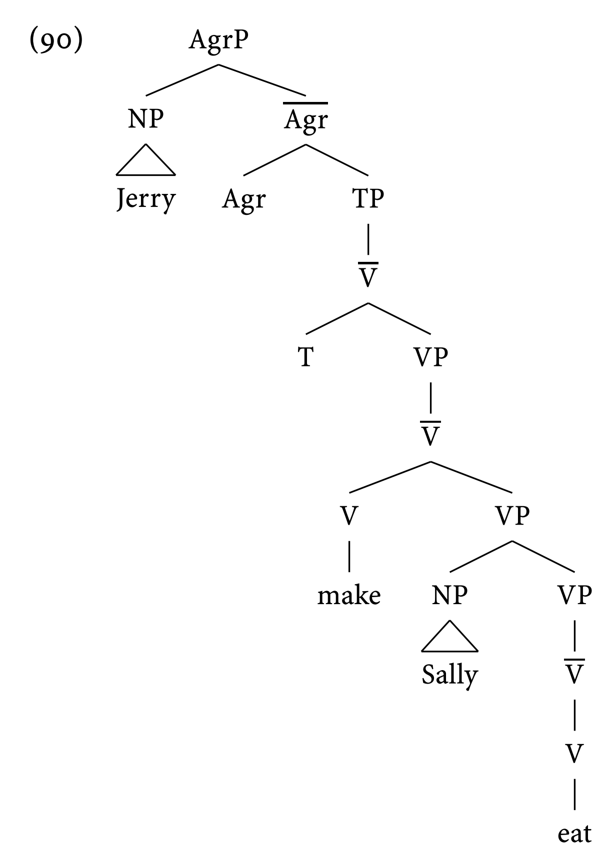
That inflectional heads assign the inflection they determine only to verbs they c-command can be appreciated by considering examples like (92) on page 199. The grammatical outcome of this structure is one in which the inflectional head determining the suffix *ing* assigns that inflection to the verb it c-commands: *ski*. What can't happen is for this I^0 to assign this inflection to *remember*, a verb it does not c-command, but is otherwise a good candidate for being inflected. (92) does not lead to the outcome in (93).

(93) * Bart should remembering Jill ski.

In general, examples of this sort are not found. If a verb bears inflection that is associated with an I^0, that I^0 c-commands the verb.

Note that this example confronts us with a question that our move to the Pollockian framework has produced. The root clause in (92) contains a modal, and this is a word that we have mapped onto the position that heads sentences. But we now have two heads, and two consequent phrases, that sentences are made from: which of these two should we map modals onto? I have decided to

(90)



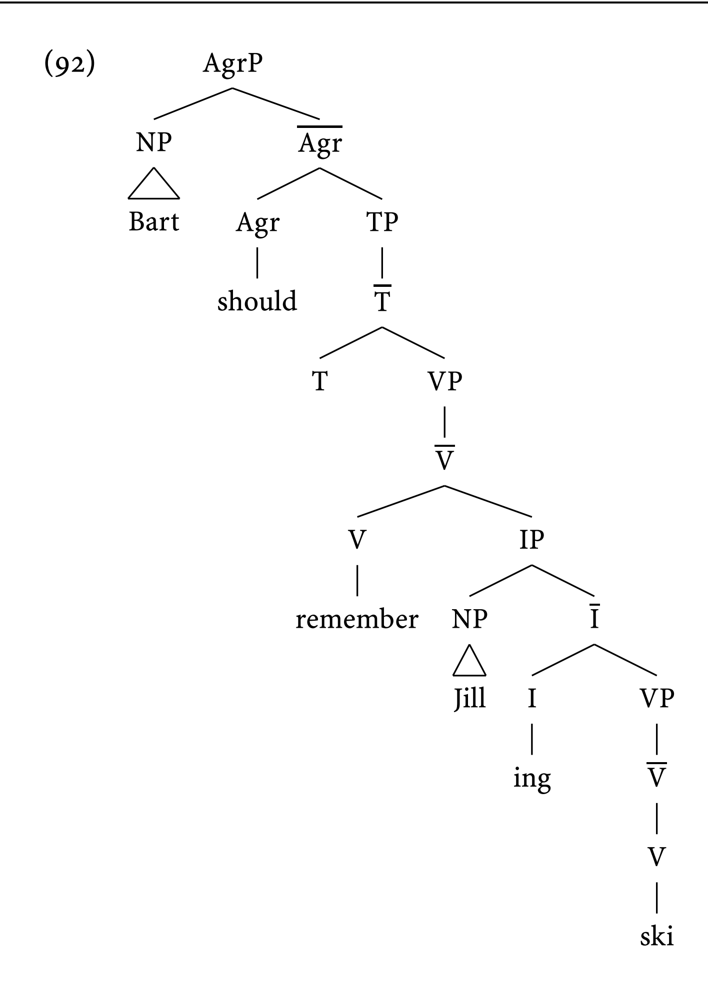
(91) * Jerry make Sally eats.

put the modal in the Agr position for two reasons. First, this correctly places modals to the left of *not*. If NegP lies between AgrP and TP, then we do not want modals to be positioned in T^0, as that would place them to the right of negation. And, second, some modals appear to have alternations which are suggestive of tense. The forms in (94) could be seen as present and past tense forms of these modals.

(94) a. *shall ~ should*
 b. *will ~ would*

While there is never any agreement morphology associated with modals in English, there might be tense morphology. It's not clear to me how to get these inflected forms from the representation in (92), but perhaps the morphology in T^0 is capable of moving to the modal in such examples.

(92)



We should endeavor to capture these parallel constraints on the ways that main and auxiliary verbs are inflected in English. It is not clear that can be done if the means by which they get their inflection is different. If, on the other hand, the process is the same in both venues, then we can place these constraints on that process and in this way explain why they are shared. That appears to be the direction presently taken.

The c-command requirement and the Head Movement Constraint are presently formulated as constraints on the movement relation. That is where they were originally discovered and that is how they were initially framed, as a result. But we can now see that there is reason to rethink this decision. Main verbs do not move, and yet the Head Movement Constraint and the c-command condition nonetheless hold between a main verb and the I^0 its inflection derives from. We want the Head Movement Constraint and the c-command conditions to hold,

therefore, on the inflection process itself. We then want those constraints to be reflected in how the inflected verbs moves to the I^0 that determines their inflection. I will sketch a proposal here that puts together some of the ideas in chapters 3 and 4 of Chomsky (1995). It is not too faithful a reproduction of those ideas, however.

First, let us formulate the condition under which inflectional heads assign verbal inflection. This relation is known as AGREE. We begin by expressing the inflectional morphology involved in terms of features. Let verbs come with features that will determine how their morphological form is determined, and let Agr^0 and T^0 also come with those features. This allows us to express the assignment of inflectional morphology in terms of matching the relevant features on a verb with the features on Agr^0 and T^0. The AGREE relation specifies under what conditions this matching of features occurs.[10]

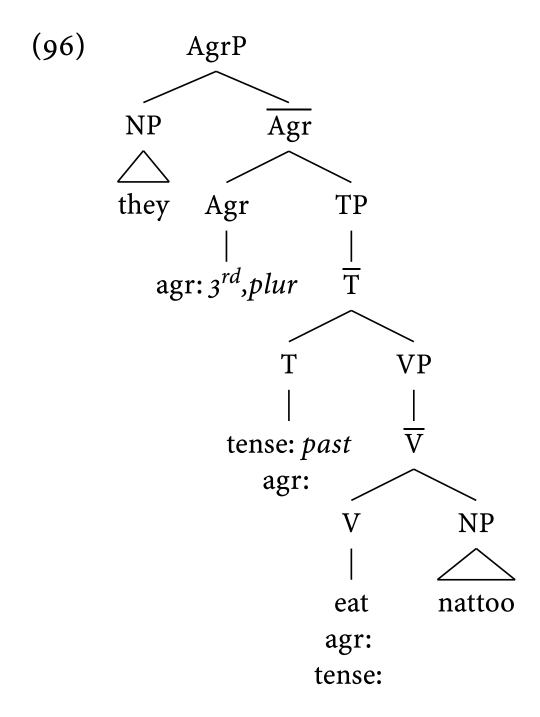
(95) Let X^0 and Y^0 be terms that carry features, and let the features on X^0 have a value and those on Y^0 be unvalued. The features on Y^0 can get the value of the matching features on X^0 just in case Y^0 AGREES with X^0. X^0 and Y^0 AGREE iff:

a. X^0 c-commands Y^0, and

b. there is no Z^0 such that X^0 c-commands Z^0 and Z^0 c-commands Y^0.

A simple transitive clause would get a representation like that in (96).

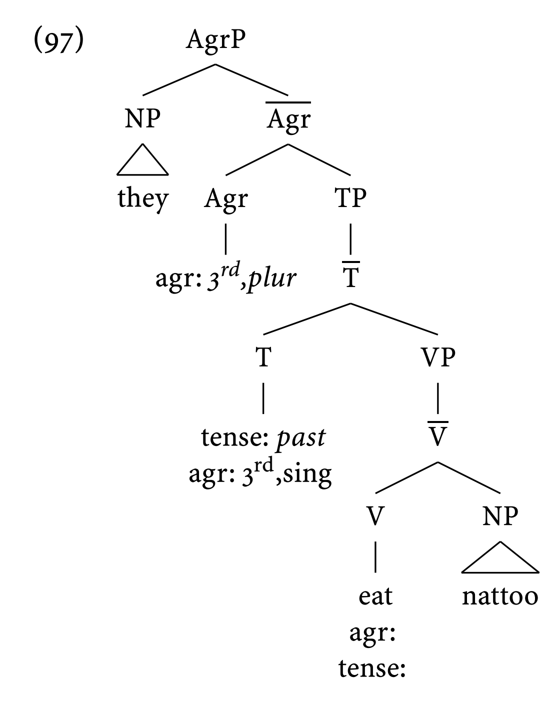
10 The system I will lay out here is modeled after Adger (2003).



The verb in this example has features for agreement and tense which, when they are given values, will let the process of lexical insertion know which lexical item to put into this position. Similarly, the T^0 position is occupied by something that has features for agreement and tense, and one of those features is also given a value, here: past. On this view, T^0 is something that contains an agreement feature; it may not be obvious why tense should be so equipped. It will be necessary to have T^0 be the place from which the agreement feature on the following verb is valued, so there is a theory internal reason for making this claim. It would be good to have an independent criterion upon which to make this claim. Perhaps the fact that the tense/agreement morpheme in English is a portmanteau could be related. Finally, Agr^0 is equipped with an agreement feature, and it has the value 3^{rd}, singular.

Because Agr^0 c-commands T^0, and there is no head between them, the value for the agreement feature on Agr^0 can value the agreement feature on T^0. This produces (97).

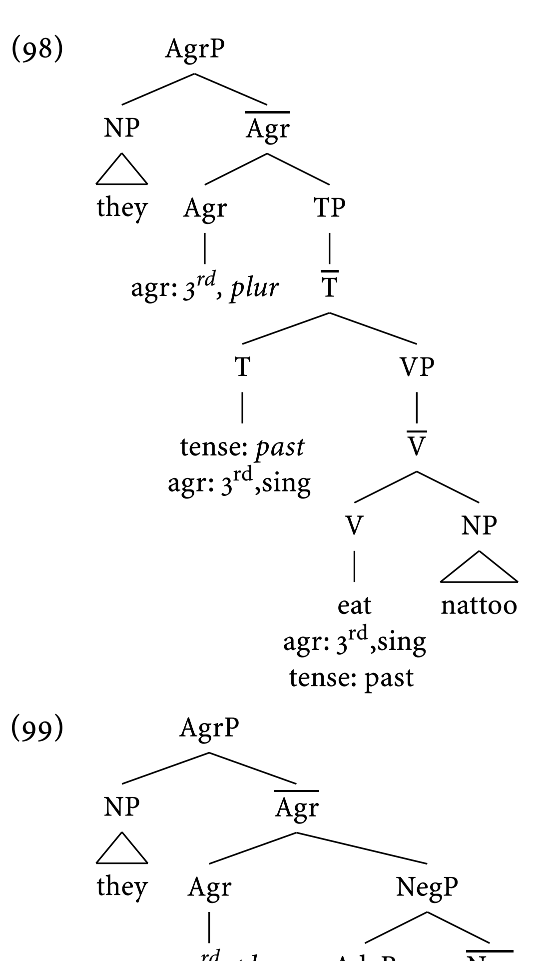
(97)



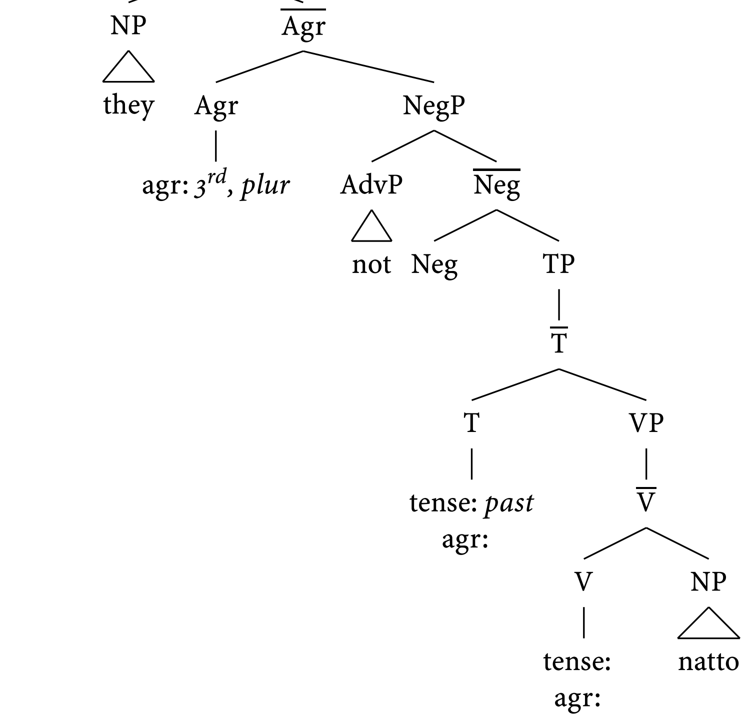
Both of the unvalued features on the verb *eat* can now be valued by the features on T^0. This is because T^0 c-commands *eat* and there is no head between them; they are therefore able to AGREE. This produces (98) on the next page. The values of inflectional features are assigned down through the heads, and verbs then express those collected feature values morphologically.

We can now model the disruptive effect that NegP has on inflecting main verbs in terms of the "Head Movement Constraint," which is now expressed as a condition on AGREE. If we assume that the head of NegP does not come with the tense or agreement features, a sentence with NegP would get a representation like (99).

(98)


AgrP
NP
Agr
they
Agr
TP
agr: 3rd, plur
T
T
VP
tense: past
agr: 3rd,sing
V
V
NP
eat
nattoo
agr: 3rd,sing
tense: past

(99)


AgrP
NP
Agr
they
Agr
NegP
agr: 3rd, plur
AdvP
Neg
not
Neg
TP
T
T
VP
tense: past
agr:
V
V
NP
tense:
agr:
natto

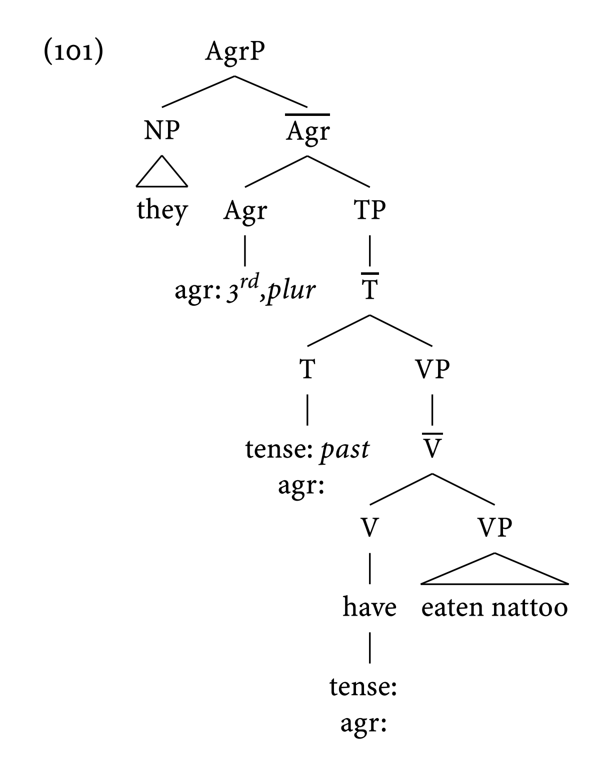
That Neg^0 does not come with agreement or tense features could be linked to the fact that negation does not express those features: English negation does not come inflected for tense or agreement. The values on the agreement feature that Agr^0 has cannot be given to the unvalued agreement feature on T^0 because Neg^0 is a head that intervenes. As a consequence, the agreement feature on the verb will also not be valued. The ungrammaticality of (99), then, can be credited to the inability of this system to inflect the verb with both tense and agreement when NegP is present. To make explicit the requirement that values of the features on T^0 and Agr^0 get expressed on a verb, let's adopt (100).

(100) The Stray Feature Filter
The values of an inflectional feature must be morphologically expressed. In English, values for tense and agreement features are expressed on verbs.

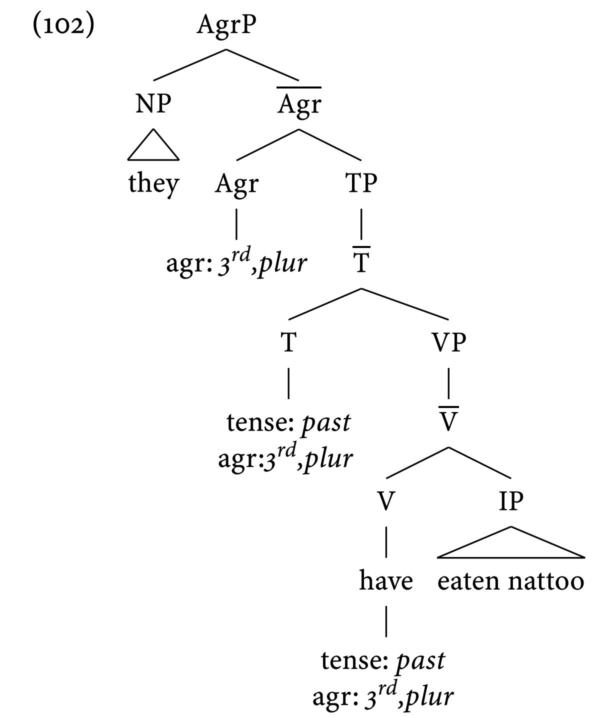
The ungrammaticality of (99), then, comes from the fact that the value for the agreement feature is not found on a term that can express that value: it is not assigned to a verb, and only verbs in English are equipped with the morphology to express agreement. The condition in (100) is just a restatement of Lasnik (1981)'s Stray Affix filter. His condition said that the bound morphology placed in an inflectional head could not be expressed without the aid of a verbal stem in that very same inflectional head. The Stray Feature Filter does nothing more than translate that into a framework which expresses inflectional morphology with features; (100) says that the inflectional feature values carried by some inflectional head must find a way of being assigned to a verbal stem. I've named it the Stray Feature Filter to remind us of this connection.

Let's now consider how this system will work with auxiliary verbs. A simple clause will have the structure in (101). The chain of AGREEment relations that we saw in the main verb case will apply here too, producing the representation in (102).

(101)


AgrP
NP
Agr
they
Agr
TP
agr: 3rd,plur
T
T
VP
tense: past
agr:
V
V
VP
have
eaten nattoo
tense:
agr:

(102)

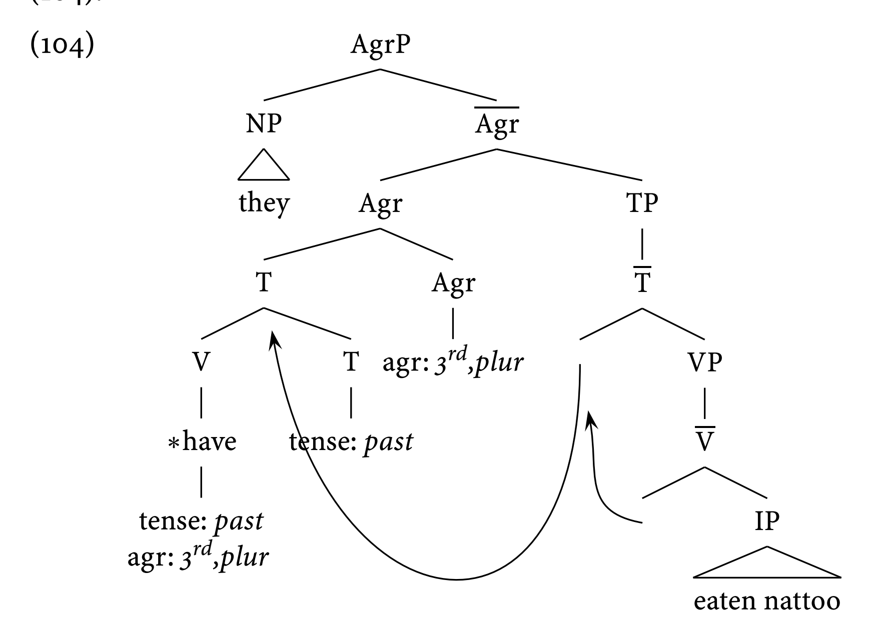

AgrP
NP
Agr
they
Agr
TP
agr: 3rd,plur
T
T
VP
tense: past
agr:3rd,plur
V
V
IP
have
eaten nattoo
tense: past
agr: 3rd,plur

This representation satisfies all of our present conditions and so should lead to a grammatical string. As we have seen, however, this isn't what we want. Auxiliary verbs surface in Agr^0 in these contexts. There must be some property that auxiliary verbs have that prevents them from surfacing in their D-structure position when they are inflected. I don't know what this property is, but it is possible, nonetheless, to characterize the consequences it has. It requires the auxiliary to adjoin to the inflectional heads which are responsible for giving values to the features the auxiliary contains. If we call this mysterious property: ∗, then we can state this requirement with (103).

(103) If a lexical item has ∗, then it must be pronounced adjoined to the heads whose features it is in an AGREE relation with.

Auxiliary verbs have ∗, and so representations like (102) will not be well-formed S-structures. Instead, the head movement operation must kick in to produce (104).

(104)



The fact that auxiliary verbs differ from main verbs in needing to move is now coded in the ∗ property. We should hope to find what the ∗ property is, and it seems likely that it has something to do with how the inflectional morphology is expressed since it is a requirement that makes reference to the features expressed by that morphology. One possibility would be to let auxiliary verbs differ from main verbs in the kinds of features they come with. If, for instance, auxiliary verbs are not equipped with the tense and agreement features, but

the morphological forms they have nonetheless express those features, then the Stray Feature Filter could not be satisfied by way of AGREEment. There will be no way of passing the values of tense and agreement onto the auxiliary verb. If the Stray Feature Filter can be satisfied by letting the syntax form a word in which the tense and agreement feature values are expressed, then we could head movement as the agent responsible for forming that word. The highest Agr^0 in (104), on this view, would satisfy the Stray Feature Filter. This interpretation of the $*$ property would amount to letting the features and the morphology that expresses those features be packaged in two different ways. In one, the features and the morphology are separated in D-structures and are brought together by syntactic means. This is what happens with auxiliary verbs in English. In the other, the morphology and the features are packaged together into one lexical item at D-structure, and the values those features acquire are determined by AGREE. This is what happens with main verbs in English.

While this seems to me a natural way of expressing what the $*$ property is in these scenarios, I'll not commit to that account in these notes. Instead, I'll continue to leave this property unaccounted for, and rely just on the descriptive consequences it has.

The $*$ property can be used to describe what it is that makes auxiliary verbs, unlike main verbs, move. But we must still explain why main verbs are unable to move. One idea that has had found some favor is to posit a force that disfavors movement. Chomsky (1995) suggests that something of this sort is responsible for blocking movement of main verbs. He observes that we could relate the fact that main verbs are not required to move, that is, they don't have the $*$ property, to the fact that they cannot move if we could build into the grammar a force that allows movement only when it is required. Chomsky's own proposal is built upon a different account of how main verbs get inflected, and so will not work for us.[11] But David Pesetsky has proposed a way of expressing this notion that is very close to Chomsky's and fits our account of verbal

11 Chomsky assumes that main verbs inflect by moving "covertly" to the inflectional heads that hold the relevant morphology. He then assigns different weights to covert and overt movement so that covert movement is favored. A difficulty with his proposals that has not yet been overcome is defining a notion of "covert" movement that can do what is required of it in other contexts. Covert movement is thought to arise in other contexts, but in those contexts it always has semantic consequences. Covert main verb movement, by contrast, does not. It appears that covert movement should be designed in such a way that it must have semantic consequences, and it is therefore not a suitable mechanism for inflecting main verbs. See Stechow (1996) and Fox (2000) for some discussion.

inflection.[12] He proposes the constraint in (105).

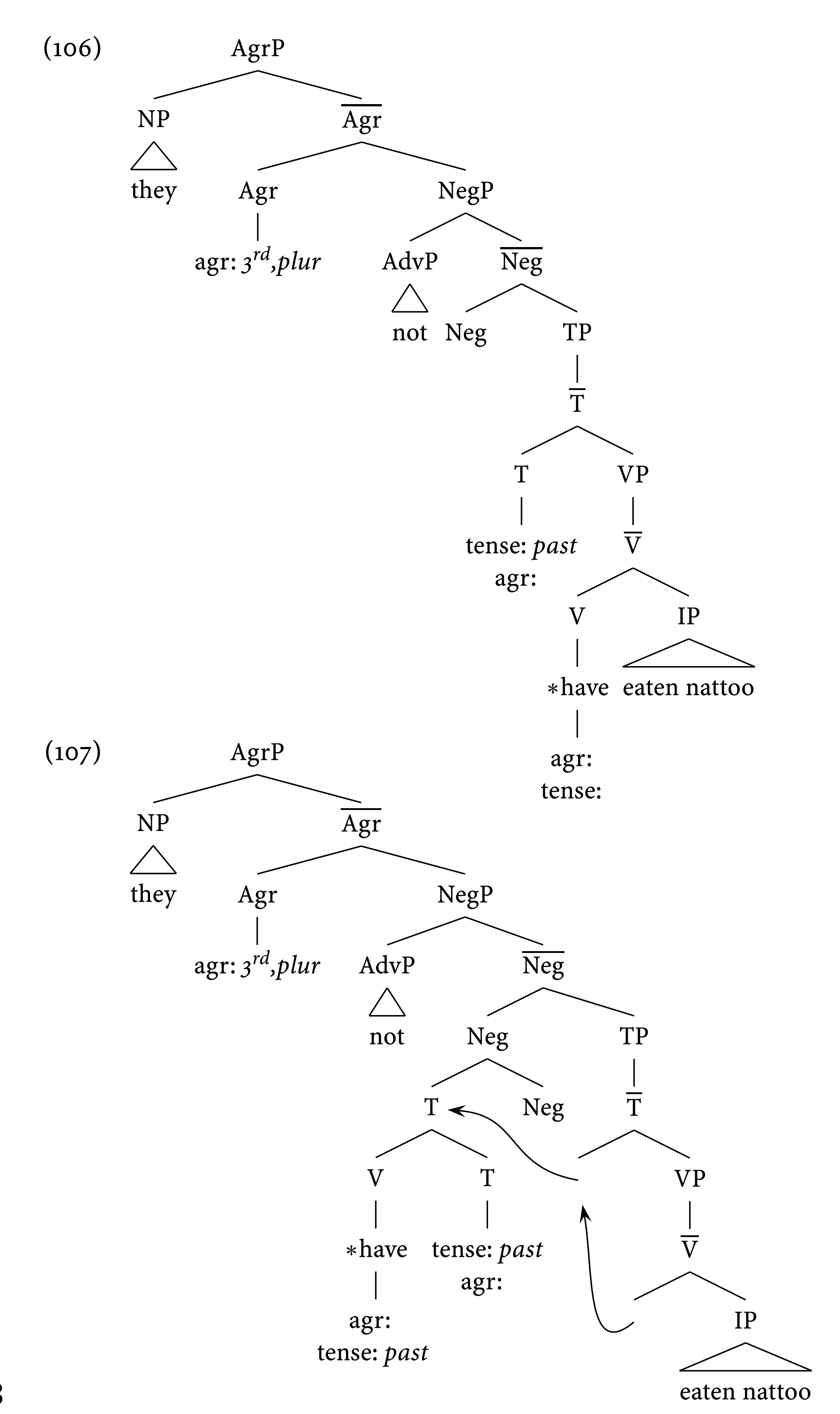
(105) Earliness
Let U be a fixed D-structure representation and $\mathscr{D} = \{D, D_1, \ldots, D_n\}$ be all the grammatical derivations from U to some representation that satisfies the conditions on well-formed S-structures. Every $D_i \in \mathscr{D}$ is ungrammatical if $|D_i|$ is greater than some other $|D|$ in $\mathscr{D}$, where "$|D|$" is the number of movement rules in D.

What Earliness requires is that derivations from D-structure to S-structure minimize the number of movement operations. If it is possible to satisfy the constraints on S-structure without moving, than movement is blocked. This is precisely what happens in the inflection of main verbs. They are able to satisfy the Stray Feature Filter by way of AGREE, and this makes movement of main verbs to T^0 or Agr^0 superfluous.

This is a popular account of the failure of main verb movement, but it is not entirely successful. This becomes apparent when the cases in which an auxiliary verb combines with NegP are considered. In such cases, we have the D-structure representation in (106) on the facing page. As we have seen in the case of main verbs, AGREE is prevented from giving the value that Agr^0 has for the agreement feature to either T^0 or the verbs beneath T^0. Just as in the case with main verbs, then, this representation does not satisfy the Stray Feature Filter. Instead, as we know, the auxiliary verb moves until it surfaces in the Agr^0 position. It seems reasonable to assume that it's the ability of the auxiliary verb to move that allows it to get inflected. We want movement and AGREE to have different outcomes in this circumstance.

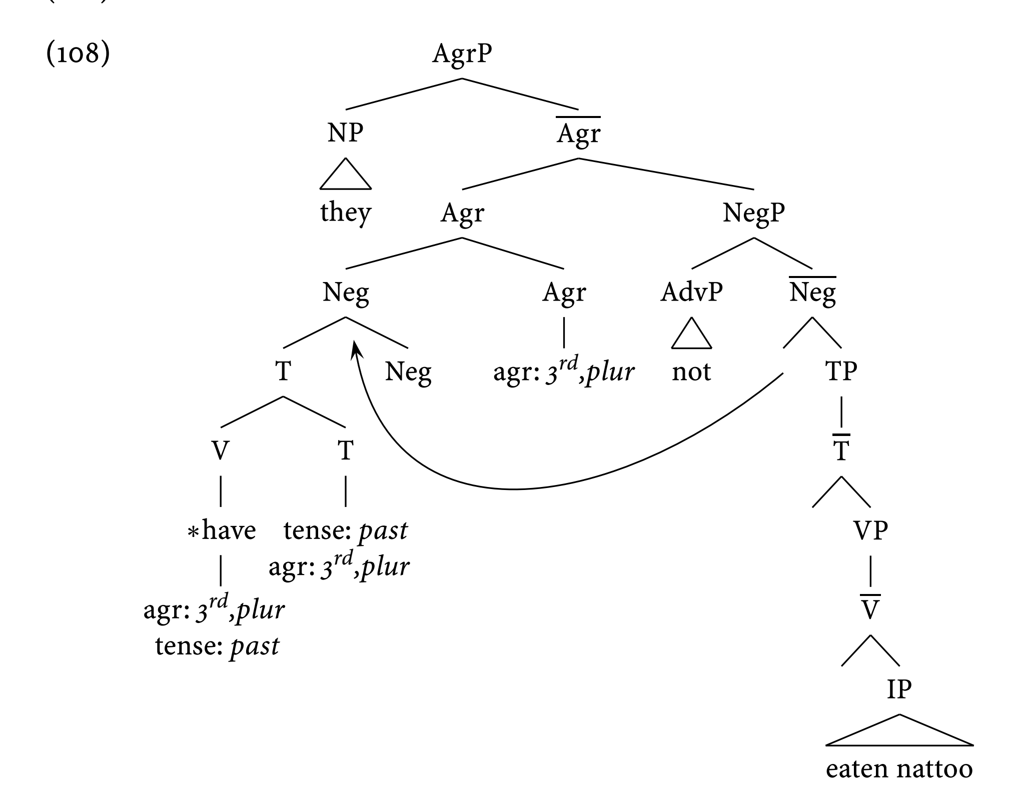
The auxiliary is close enough to T^0 to get its tense feature valued. If it moves to T^0, and the resulting complex moves to Neg^0, then the representation in (107) will be produced.

12 The proposal comes from the unpublished Pesetsky (1989), which can be found at his website. This is not his formulation of the condition, but I think it expresses his idea.


(106)
AgrP
NP
they
Agr
Agr
agr: 3rd,plur
NegP
AdvP
not
Neg
Neg
TP
T
T
tense: past
agr:
VP
V
V
∗have
agr:
tense:
IP
eaten nattoo
(107)
AgrP
NP
they
Agr
Agr
agr: 3rd,plur
NegP
AdvP
not
Neg
Neg
T
V
∗have
agr:
tense: past
T
tense: past
agr:
Neg
TP
T
VP
V
IP
eaten nattoo

In this representation, *have*, and the T^0 it has joined with, are both part of Neg^0. Perhaps be virtue of being part of Neg^0, *have* and T^0 are now no longer separated from Agr^0 by another head. If so, the agreement features they contain can be valued by Agr^0. This would satisfy the Stray Feature Filter. In order to satisfy the ∗ property, head movement would move Neg^0 into Agr^0, producing (108) below. This is the desired outcome.

(108)



But if this is how auxiliary verbs manage to overcome the disruptive effect that NegP has on getting inflected, then Earliness will not be sufficient to prevent main verbs from making use of the same method of getting inflected when NegP is present. In fact, the point is more general than the particular proposal I've made here. Irrespective of details, if it's movement that allows auxiliary verbs to get inflected when NegP is present, then Earliness will not prevent main verbs from moving when NegP is present. Earliness will prevent main verbs from moving when there is no NegP. But it cannot block derivations that are otherwise legitimate if they are the only ones that give rise to a well-formed S-structure; and that is precisely the case for a derivation which would move a main verb through Neg^0 to Agr^0. Something must be added to

this system if it is to block main verbs from moving when NegP is present. There are proposals in the literature, but none which have gained consensus acceptance. This remains an open problem, I believe. For this proposal to be complete, then, something must be found that derives (109).

(109) Main verbs cannot combine with Neg^0.

This will block derivations in which main verbs get close enough to Agr^0 to have their agreement feature valued and thereby remove the only derivation in which can satisfy the Stray Feature filter when a main verb is separated from Agr^0 by Neg^0.

Earliness also faces problems from examples in which head movement appears to be optional. We've seen something with that appearance in two contexts. In French infinitival clauses, we've seen that auxiliary verbs are able to surface in one of two positions: to the left of *pas* or to the right of *pas*. This is illustrated by (110), repeated here.

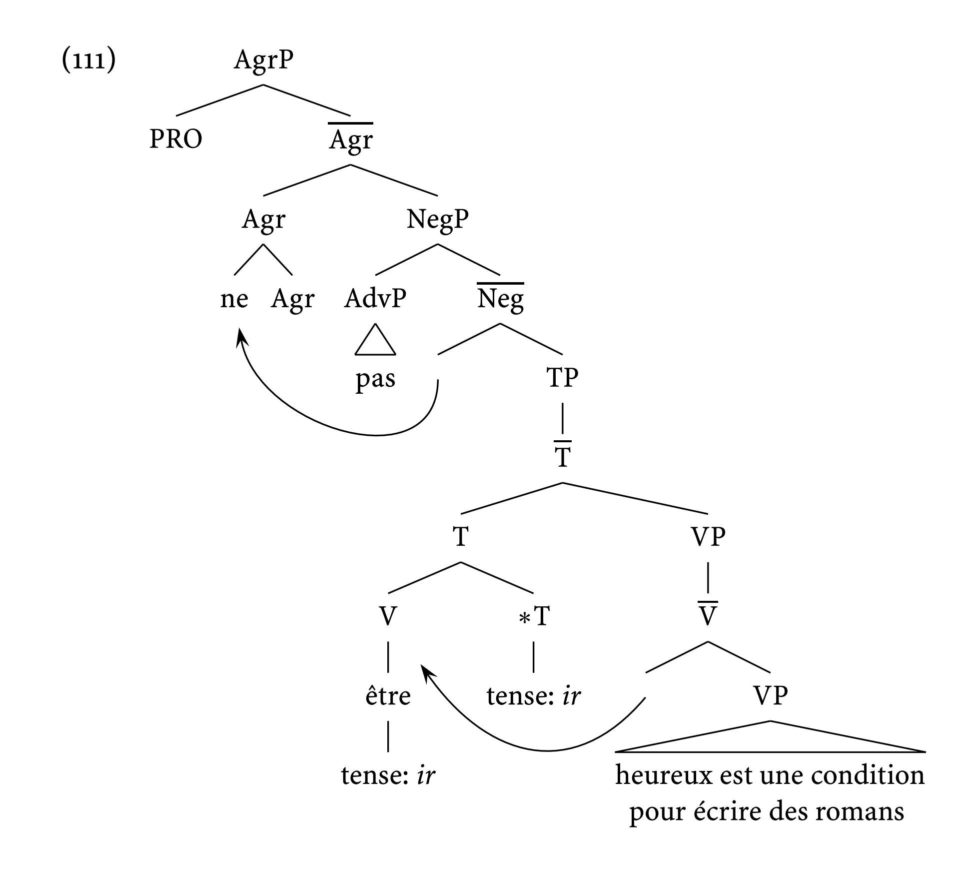
(110) a. ne *pas* **être** heureux est une condition pour écrire des romans
ne not be happy is a condition for writing novels
'To not be happy is a condition for writing novels.'

b. N'**être** *pas* heureux est une condition pour écrire des romans.
ne'be not happy is a condition for writing novels
'To not be happy is a condition for writing novels.'

Consider the representation given to (110a), in (111) on the following page. I have given T^0 the ∗ property, as this head forces head movement of both auxiliary and main verbs to it. (Recall: we saw that main verbs in French infinitival clauses move to T^0 and are thereby separated from their NP complements by adverbs.) In general, we can characterize those inflectional heads that allow main and auxiliary verbs to move to them, i.e., those heads that are described as being "strong," as having the ∗ property.

Consider now the status of Agr^0 in this representation. I suggested in the previous section that this head is not associated with inflectional morphology. Such heads, I suggested, could serve as the site of head movement but did not force head movement. For this reason, auxiliary verbs could optionally move to Agr^0 in infinitival clauses, and when this option is taken, (110b) arises. But this account is incompatible with Earliness, which would block the needless movement required to derive the word order in (110b). Let's consider how we might characterize these cases in a system of inflection assignment that includes Earliness.

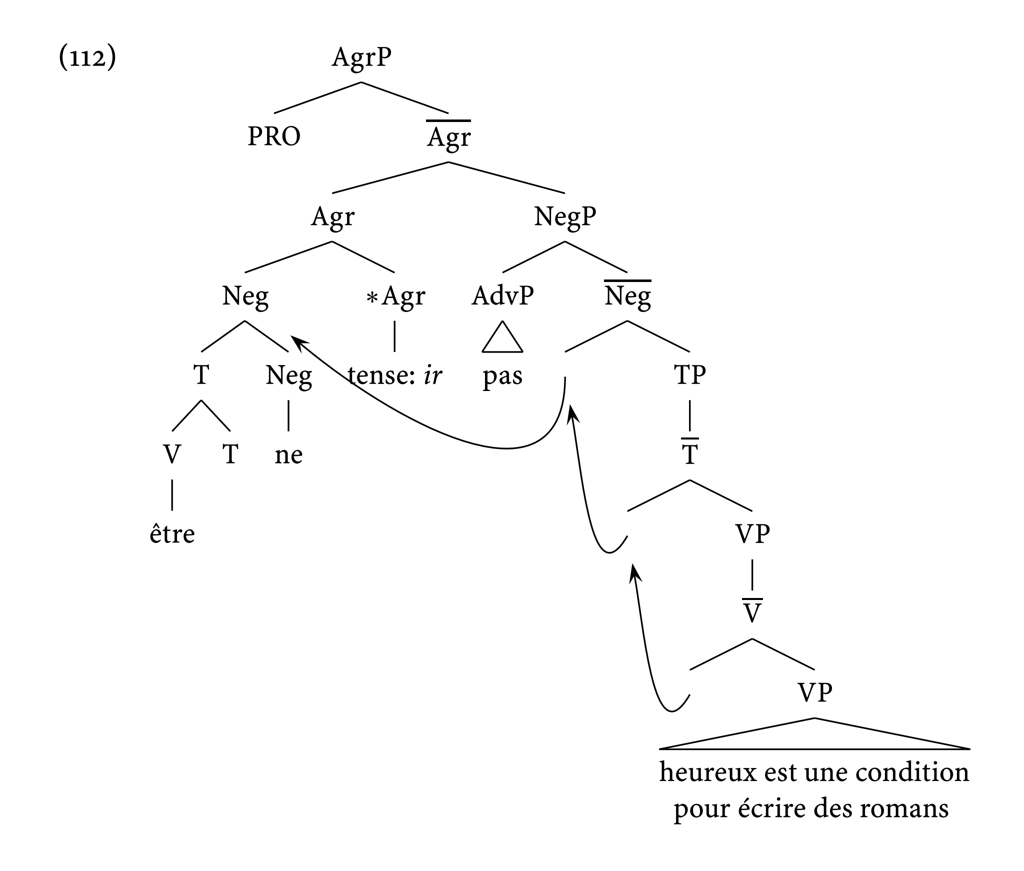
(111)



First: if Agr^0 has feature values that need to be assigned, then we should expect the presence of NegP to block assigning those values to heads beneath Neg^0. Therefore, when no verb moves to Agr^0, the present system requires that Agr^0 be associated with no verbal morphology. And, as noted, in that case movement to Agr^0 will be prevented by Earliness. The sentence in (110a) is therefore consistent with our present system under the assumption that Agr^0 has no feature values to assign.

But the outcome in (110b) requires a different set of assumptions about the status of Agr^0. In that circumstance, Agr^0 must be associated with verbal morphology. Because the inflectional morphology in these two scenarios remains the same, let's assume that in the cases where a verb moves to Agr^0, infinitival morphology is associated with Agr^0 and not T^0. Under the assumption that the $*$ property is associated with particular morphologies, we might speculate that when the infinitival morphology resides in Agr^0 it gives Agr^0 the $*$ property. A way of characterizing (110b) in the theory of this section gives it the

(112)



representation in (112) below.[13] The ∗ property requires that a verb expressing the value the tense feature be adjoined to Agr^0, and this results in the auxiliary verb moving through Neg^0 to Agr^0. Since this is the only derivation in which the ∗ property gets its requirements satisfied, Earliness permits it. In these situations, when there is no auxiliary verb, and instead a main verb expresses the value of the tense feature, there will be no grammatical outcome. This is because main verbs, as we've seen, are incapable of moving through Neg^0.

To characterize "optional movement," then, a system which embraces Earliness requires that there be two possible representations: one that forces movement and another in which it is blocked. Under the particulars of the account I am sketching here, that would mean that French speakers who allow the alternation in (110) allow the valued tense feature to be generated in either Agr^0

13 Note that this representation does not linearize *ne* relative to *être* correctly. We will continue to abstract away from the issue of how to match a linearized form to the complexes created by Head Movement.

or T^0. When it is mapped onto T^0, main and auxiliary verbs will be forced to move into T^0 because of the requirements of the $*$ property associated with the tense feature. When it is mapped onto Agr^0, auxiliary verbs will be forced to move to Agr^0, again because of the requirements of the $*$ property associated with the tense feature, and main verbs will simply be ungrammatical. The only grammatical outcome for infinitival clauses with just main verbs, then, will be ones in which the valued tense feature is mapped onto T^0.

We could extend this reasoning to the other situation where we've conjectured that there is optional movement. This situation arises in English infinitival clauses and responds to the fact that the infinitival marker *to* is able to either follow or precede *not*.

(113) a. Maribel tried not to eat nattoo.

b. Maribel tried to not eat nattoo.

Rather than seeing (113b) as being derived by movement from (113a), we could see this alternation as a reflection of the ambiguous status of *to*. Perhaps English speakers allow *to* to stand in either the T^0 or Agr^0 positions; this would give the two word-orders in (113).

If these methods for characterizing the flexible position of certain heads can be confirmed, then we can continue to entertain the hypothesis that Earliness exists, and rely on its consequences to block main verbs from moving in English (non-negated) sentences. In any case, let's assume that verbs inflect by way of the AGREE relation, and that the Head Movement Constraint is a constraint on this relation and not movement. As we've seen, in order to spread this constraint on AGREE to cases where head movement arises, we need a method of explaining how head movement is tied to AGREE. Head movement needs to be seen as one outcome of the AGREE relation, and therefore an outcome that inherits the constraints on AGREE. I've sketched a way of doing that in this section, but it's clear that this is not yet a full account. It relies on finding a source for the $*$ property, and for understanding what prevents main verbs from moving through Neg^0 into higher heads.

4.6 AGREE and Case

We have now encountered two systems in which inflectional morphology is assigned by one term to another. In addition to the cases examined in this chapter, where an inflectional head assigns feature values to verbs that they c-command, we made use of a system of Case assignment in the previous chapter

to determine the position of argument NPs. These two systems do something very similar: they both determine the inflectional morphology on a term by way of the proximity that term has to an item whose presence seems to be responsible for giving a value to that inflectional morphology. In the scenarios involving Case, the values range over nominative, accusative and genitive; and the terms that seem to be responsible for fixing those values are finite (and maybe some infinitival) Agr^0 (for nominative), certain verbs and prepositions (for accusative), and something we've not yet identified for genitive. Given the similarities in what these two systems are doing, it is reasonable to expect that they are the same system. We'll consider in this section how the two might be collapsed.

I'll continue to set aside how genitive Case is assigned, and restrict our attention to just nominative and accusative. Presently, the two systems are as defined in (114) and (115).

(114) Verbal Inflection

a. Let X^0 and Y^0 be terms that carry features, and let the features on X^0 have a value and those on Y^0 be unvalued. The features on Y^0 can get the value of the matching features on X^0 just in case Y^0 AGREES with X^0. X^0 and Y^0 AGREE iff:
   i. X^0 c-commands Y^0, and
   ii. there is no Z^0 such that X^0 c-commands Z^0 and Z^0 c-commands Y^0.

b. The Stray Feature Filter
The values of an inflectional feature must be morphologically expressed.

c. If a lexical item has $*$, then it must be pronounced adjoined to the heads whose features it is in an AGREE relation with.

d. Feature assignments
   i. An agreement feature is part of certain Agr^0s, T^0s and verbs. It comes with a value on some Agr^0s.
   ii. A tense feature is part of certain T^0s and verbs. It comes valued on some T^0s.

(115) Case Assignment

a. Nominative Case is assigned by finite Agr^0 to the Specifier of the phrase it heads.

b. Accusative Case is assigned by some verbs and prepositions to positions they are:

i. c-command, and
ii. adjacent to, and
iii. klose to.

c. α is klose to β if no more than one phrase dominates β but not α.

d. The Case Filter
An NP must be in a Case-marked position at S-structure.

To bring these two together, I'll start by translating Case into the feature based system that we've expressed verbal inflection with. We'll let certain Agr^0s, verbs and prepositions come with a value for the Case feature, and we'll assume that every NP has an unvalued Case feature. We'll model Case assignment, then, as fixing values for the Case feature under AGREE. Taking this step will create the system in (116)–(120d).

(116) Let X and Y be terms that carry features, and let the features on X have a value and those on Y be unvalued. The features on Y can get the value of the matching features on X just in case Y AGREEs with X. X and Y AGREE iff:

a. X c-commands Y, and
b. X is klose to Y.

(117) X is klose to Y iff:

a. For Y a head, there is no Z^0 such that X c-commands Z^0 and Z^0 c-commands Y^0.
b. For Y a phrase, there is no more than phrase that dominates Y and not X.

(118) Feature valuation Requirement (holds of S-structure)
The values of an inflectional feature must be morphologically expressed and every feature must get a value.

(119) If a lexical item has $*$, then it must be pronounced adjoined to the heads whose features it is in an AGREE relation with.

(120) Feature assignments

a. An agreement feature is part of certain Agr^0s, T^0s and verbs. It comes with a value on some Agr^0s.
b. A tense feature is part of certain T^0s and verbs. It comes valued on some T^0s.
c. A unvalued Case feature is part of every argument NP.

d. A valued Case feature is part of certain Agr^0s, verbs and prepositions.

The following changes have been made. First, the terms that can enter into an AGREE relation have been widened to include phrases as well as heads. The locality condition imposed on the term that has the value to assign and the term whose feature gets assigned that value has been relativized so that it discriminates between the cases in which a head or phrase are involved. If a value is assigned to a head, then the Head Movement Constraint applies; if a value is assigned to a phrase, then the definition of klose used for Case assignment is employed.

Next, the Case Filter has been stated as a condition on features. The requirement is now that unvalued features must get a value before the term that bears those features can be pronounced. This condition has been added to the Stray Feature filter, which requires that values get assigned to terms that can express them. The combined condition is renamed the Feature Valuation Requirement. Finally, valued Case features are stated to be part of our former roster of Case assigners, and unvalued Case features are stated to be part of the NPs that the Case filter holds of.

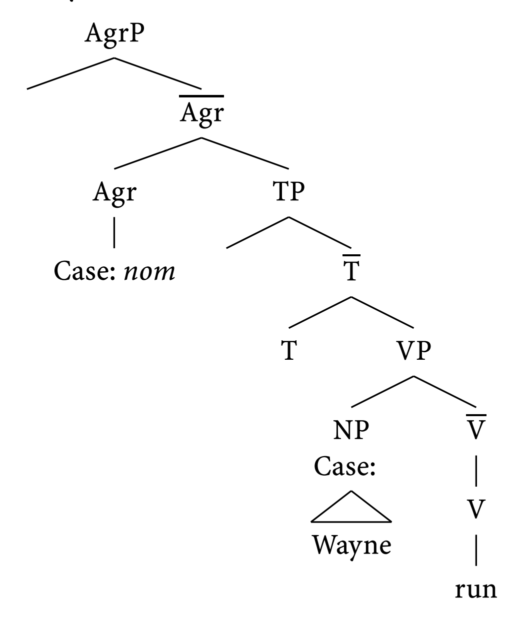
This reformulation leaves out the condition under which nominative Case assignment was formerly thought to apply. That condition had nominative Case assigned to the Specifier position of AgrP by certain Agr^0s. It seems that it is the Agr^0 that determines nominative Case because it appears that whether nominative Case is available depends on the kind of clause involved, and the "kinds" can be distinguished by their heads. The assumption that nominative Case is assigned to the Specifier of AgrPs, is assumed because it allows for an explanation of why argument NPs seem to be forced to move into that position when no other Case is available to them. Together, these two assumptions require that nominative Case be assigned by Agr^0 to their Specifier positions. To fold nominative Case assignment into this system requires we abandon one or the other of these assumptions.

In fact, both assumptions are suspect. As you might have noticed, the fact that Argument NPs must move to Specifier of AgrP is often redundantly captured by the EPP. We might conjecture, then, that nominative Case is assigned to some other position, and let A movement to Specifier of AgrP be driven by the need to satisfy the EPP. On the other hand, we might be tempted to assume that it is Agr^0 that assigns nominative Case. In fact, if we look at the clauses whose Specifier positions receive nominative Case, they are all, except

perhaps for root clauses, ones that are embedded under a complementizer. We might imagine that it is the presence of the complementizer that correlates with the availability of nominative Case, and from this we could conclude that nominative Case is assigned by complementizers. Both of these directions have been pursued. I will sketch a way of pursuing the first: one that maintains the assumption that nominative Case comes from Agr^0s, but assigns it under c-command.

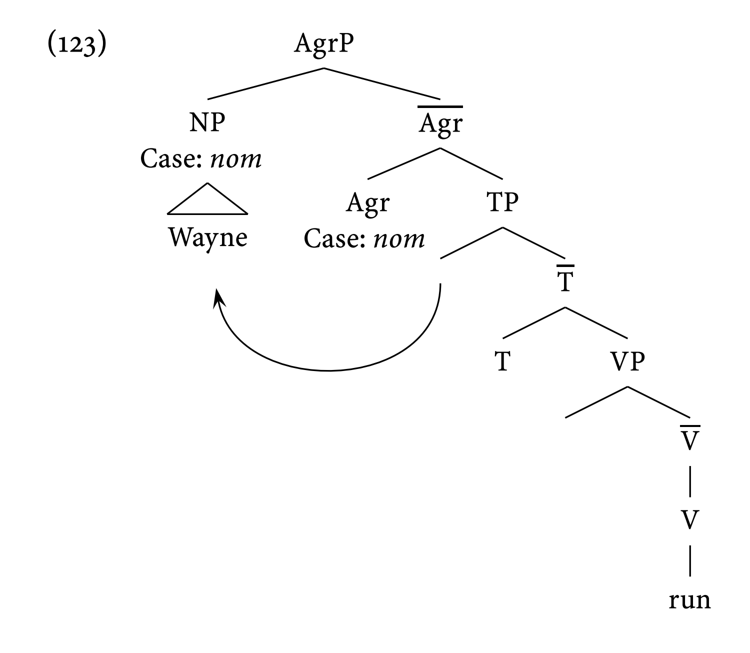
Consider what would be expected if finite Agr^0 assigns Case under the conditions that prepositions and verbs do. In examples in which no NegP is present, we would expect the Specifier of TP to be the position that nominative Case is assigned to. A sentence like (121) would have the D-structure indicated.

(121) Wayne ran.

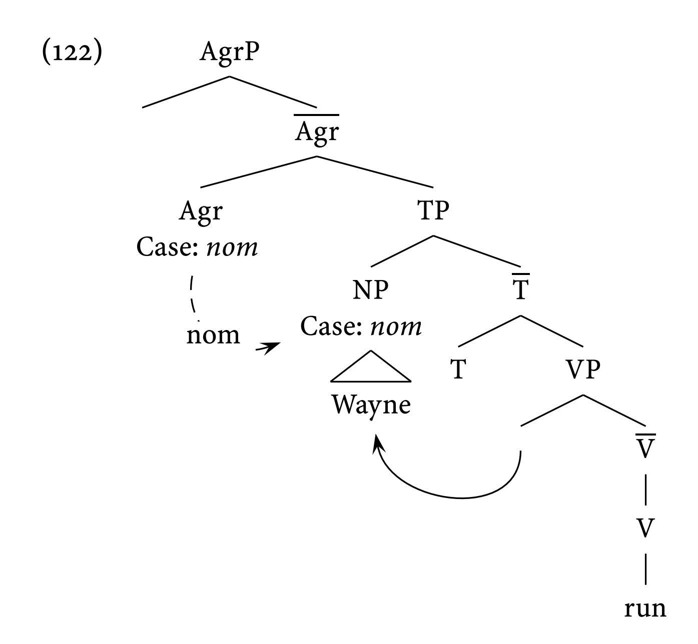


(Only the Case features are indicated in (121).) The NP bears an unvalued Case feature, a feature that must be valued before the Feature Valuation Requirement will be satisfied. In its D-structure position, it is too far from Agr^0 for AGREE to hold, and so it cannot get its value from Agr^0. If it moves to Specifier of TP, however, it will then be klose enough for AGREE to hold, as (122) shows.

(123)



(122)



This representation satisfies the Feature Valuation Requirement (the former Case Filter), but it does not yet satisfy the EPP, which requires further movement of *Wayne* to Specifier of AgrP. The result is (123) below, which corresponds to the grammatical S-structure outcome for this sentence.

The first thing we can observe, then, is that under this interpretation of how nominative Case is assigned, Argument movement must be allowed to move arguments from positions in which they get Case marked. We have seen cases which suggest Argument movement should be disallowed from moving from Case marked positions; (124) is one such example.

(124) * Wayne seemed to that pigs fly.

Adopting this view of nominative Case assignment therefore requires resolving this conflict.

One proposal is to allow movement from Case marked positions only when the argument moves to the Specifier of the head responsible for assigning that Case. Indeed, one proposal is to reduce the EPP to the $*$ property. Imagine that the nominative feature associated with Agr^0 gives Agr^0 the $*$ property. The $*$ property would then require that Agr^0 be adjoined to the NP whose Case feature it values. The Likes-Attracts-Likes Constraint, however, prevents phrases from adjoining to heads. But it will allow a phrase to move into the Specifier of the phrase that a head projects. If we loosen the statement of the requirement that the $*$ property invokes so that the Likes-Attracts-Likes Constraint is left to determine how the moving term adjoins to the head with the $*$ property, then we could see the EPP as just another instance of the $*$ property. The requirement invoked by the $*$ property could be expressed as (125).

(125) If a lexical item, α, has $*$, then it must be pronounced with the term whose features it is in an AGREE relation with adjoined to a projection of α.

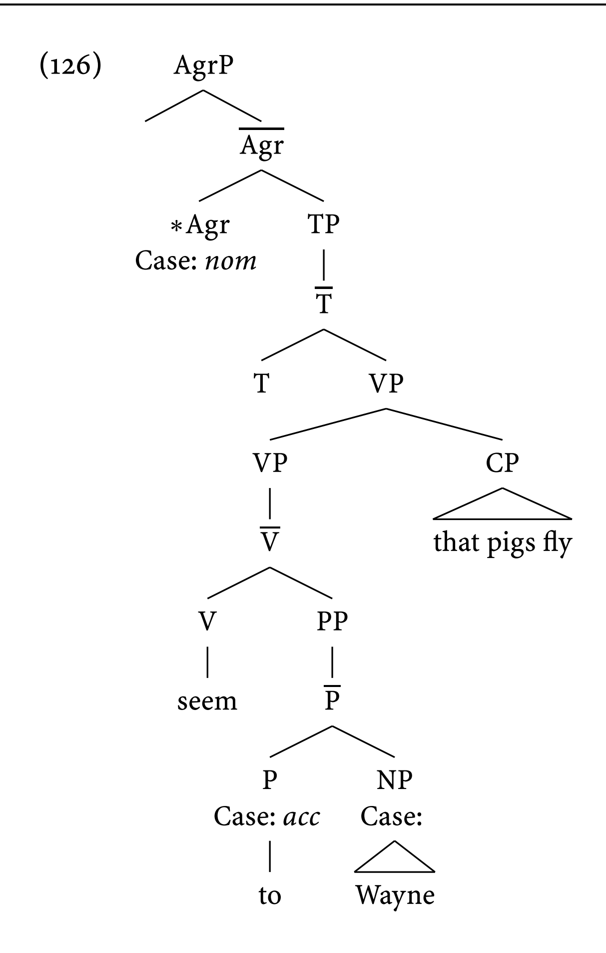
This will explain why Argument movement is blocked in examples like (124), but allowed in (123). Consider (126) on the next page: the D-structure representation (124) gets. The Case feature associated with *Wayne* can get a value only from the preposition *to*, as this is the only term with a valued Case feature that is klose enough to it. This means that Agr^0 cannot assign its Case feature value to this, or any other, NP. As a consequence, the $*$ property on Agr^0 cannot be satisfied.

4.6.1 Low Subjects

English

This reworking of the conditions under which nominative Case assignment happens claims that nominative Case is assigned to a position lower than where

(126)



the NP bearing that Case surfaces. There are examples which can be construed as evidence for this claim. In these examples, the nominative Case marked NP surfaces in a position c-commanded by the Case-assigning Agr^0, and the EPP is satisfied with *there*. One of these examples is (127).

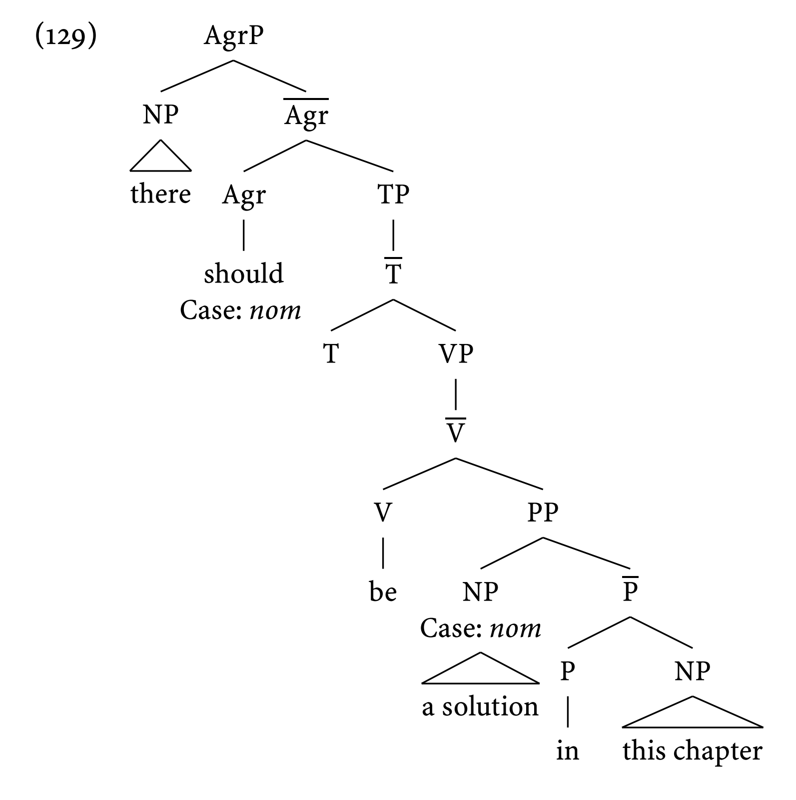
(127) There should be a solution in this chapter.

It doesn't seem that *there* in this example has the same meaning that the locative *there* has in (128).

(128) A solution stands there.

Indeed, the distribution of the *there* that is found in (127) is what is expected if it is an expletive. Just as in the scenarios in which the expletive *it* appears, we might conjecture that the EPP (or * property) is satisfied by virtue of *there*

standing in Specifier of AgrP. We know from other examples that the copular verb *be* does not assign accusative Case, so we can also conjecture that the Case which *a solution* bears is nominative. Under these assumptions, an S-structure parse for (127) might be (129) below.

(129)


If this is the right interpretation of these examples, then it confirms the hypothesis that nominative Case assignment occurs under c-command, just as accusative Case assignment does. But it presents a host of problems for the particular proposals we have arrived at. If *there* satisfies the EPP in this example, then collapsing the EPP and the ∗ property becomes difficult. Further, the distance between Agr^0 and the NP it Case marks in (129) does not fit to our definition of klose: there is clearly more than one phrase that dominates *a solution* but not Agr^0. To fit this analysis to our system would require a different definition of klose, then.

Moreover, examples of this type are governed by a set of constraints that remain little understood. The semantic type of the nominative Case marked NP matters, for instance. That argument must be an indefinite of a certain type, as the contrasts in (130) illustrate.

(130) a. There shouldn't be any solutions in this chapter.
b. There should be solutions in this chapter.
c. There should be no solutions in this chapter.
d. There should be many solutions in this chapter.
e. * There should be every solution in this chapter.
f. * There should be the solution in this chapter.
g. * There should be most solutions in this chapter.
h. * There should be all solutions in this chapter.
i. * There should be both solutions in this chapter.
j. * There should be each solution in this chapter.
k. * There should be your solution in this chapter.

And the kind of predicate involved also matters. While PPs support this construction, APs and NPs do not.

(131) a. * There is a man a fink.
compare: A man is a fink.
b. * There is a man happy.
compare: A man is happy.

VPs support the construction, but for many verbs, only when they are embedded under *be*.

(132) a. There was a woman given a book.
b. There was a woman drinking milk.
c. * There has a woman given me a book.
d. * There has a woman drunk milk.
e. * There a woman drank milk.

The exception are some (but not all) unaccusative verbs, as in (133).

(133) a. There arrived a woman.
b. There appeared an elf.

And in these cases, as can be seen, the nominative Case marked NP remains in its θ-marked object position.[14]

14 For some discussion of the constraints on the semantic type of the NP, see Milsark (1974), Safir (1985, 1987), Heim (1987) and Diesing (1992). For discussion of the constraints on the

A better understanding of how to assimilate this construction to the general theory of Case assignment will probably appear once the source of these constraints is found.

Subject Inversion Languages

There are languages in which Nominative Case-marked NPs surface in S-structure positions that, like those we've just seen in English, are c-commanded by Agr^0. In these languages, interestingly, there is no expletive like *there* involved. Spanish, Italian and Greek are such languages. Examples from Spanish and Greek are in (134).

(134) a. Leyo Juan el libro.
read John the book
'John read the book.'

b. Pandreftike o Petros tin Ilektra.
married Peter Ilektra
'Peter married Electra.'
(Alexiadou and Anagnostopoulou 1998, (3c,d): 492)

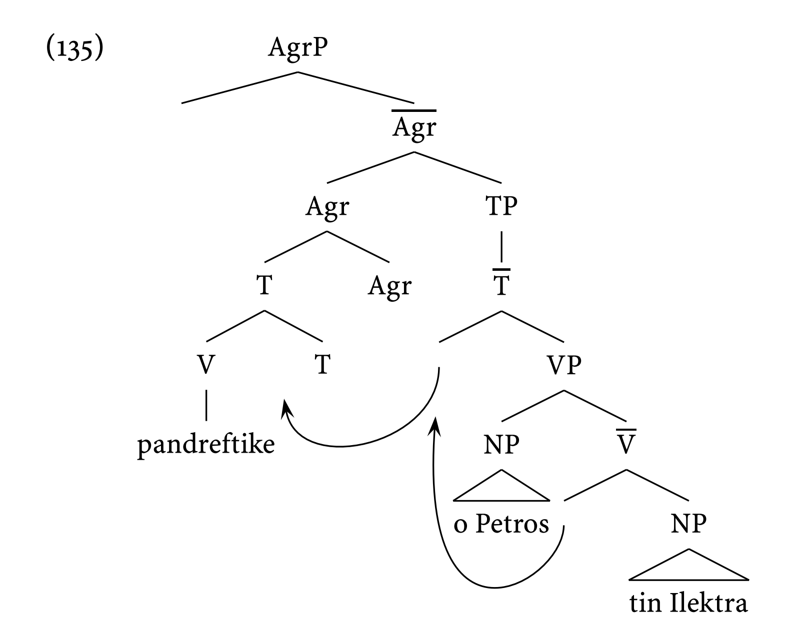
We might venture a parse for (134b) like that in (135) on the next page. These examples pose the same problem for our definition of klose that the ones in English did, but in these cases there appears to be no problem for expressing the EPP with the $*$ property. That no expletive is required in the Specifier of AgrP in these examples suggests that these are languages that have no EPP.

Note that in (135), the main verb has moved to Agr^0, thereby putting it in a position that the linearization principles of Greek will cause it to precede the nominative Case marked NP. This is a very common trait of languages that allow their nominative NPs to remain low like Greek that those NPs follow the verb (when they are not moved by other processes like wh-movement). In particular, the language properties in (136) seem to hold.

(136) If a language has (136a)–(136b), then its nominative NPs will follow the finite verb.

predicates, see Levin (1993). There is a large, and heterogeneous, literature that addresses the problem of assimilating the way *there* satisfies the EPP to a more general theory of the EPP. Among many others, see Chomsky (1995), Lasnik (1992, 1995), Moro (1997), Chomsky (1998, 2000, 2001), Bošković (2002b,a), Groat (1999, 1995), Hazout (2004), and Epstein, Pires, and Seely (2005). A good introduction to the language variation in expletives is Vikner (1995).

(135)



a. Allows nominative NPs to surface c-commanded by Agr^0.

b. Linearizes VPs and the heads that verbs move to so that they precede their complements

c. Has no expletive

These languages are said to display "subject inversion." Perlmutter (1971) and Rizzi (1982) proposed that subject inversion is always found in languages that don't have the EPP, and are verb initial. One influential attempt to explain this correlation is in Alexiadou and Anagnostopoulou (1998, 1999). They argue that, in fact, the EPP does hold in these languages, but that it is satisfied by moving the verb instead of the nominative Case-marked NP.

For that to make sense under an interpretation of the EPP that collapses it with the $*$ property, we cannot express the EPP as a property of Case. Verbs do not carry morphology that expresses Case features, and so they will not form an AGREE relation with the Case feature associated with Agr^0. If the EPP is the $*$ attribute associated with Agr^0's Case feature, then our present formulation of what the $*$ attribute requires would not be satisfied by moving the verb to Agr^0. For the proposal in Alexiadou and Anagnostopoulou (1998) to make sense, we will have to see the $*$ attribute as being associated with some feature that both the verb and the nominative Case marked NP express. That feature, as it turns

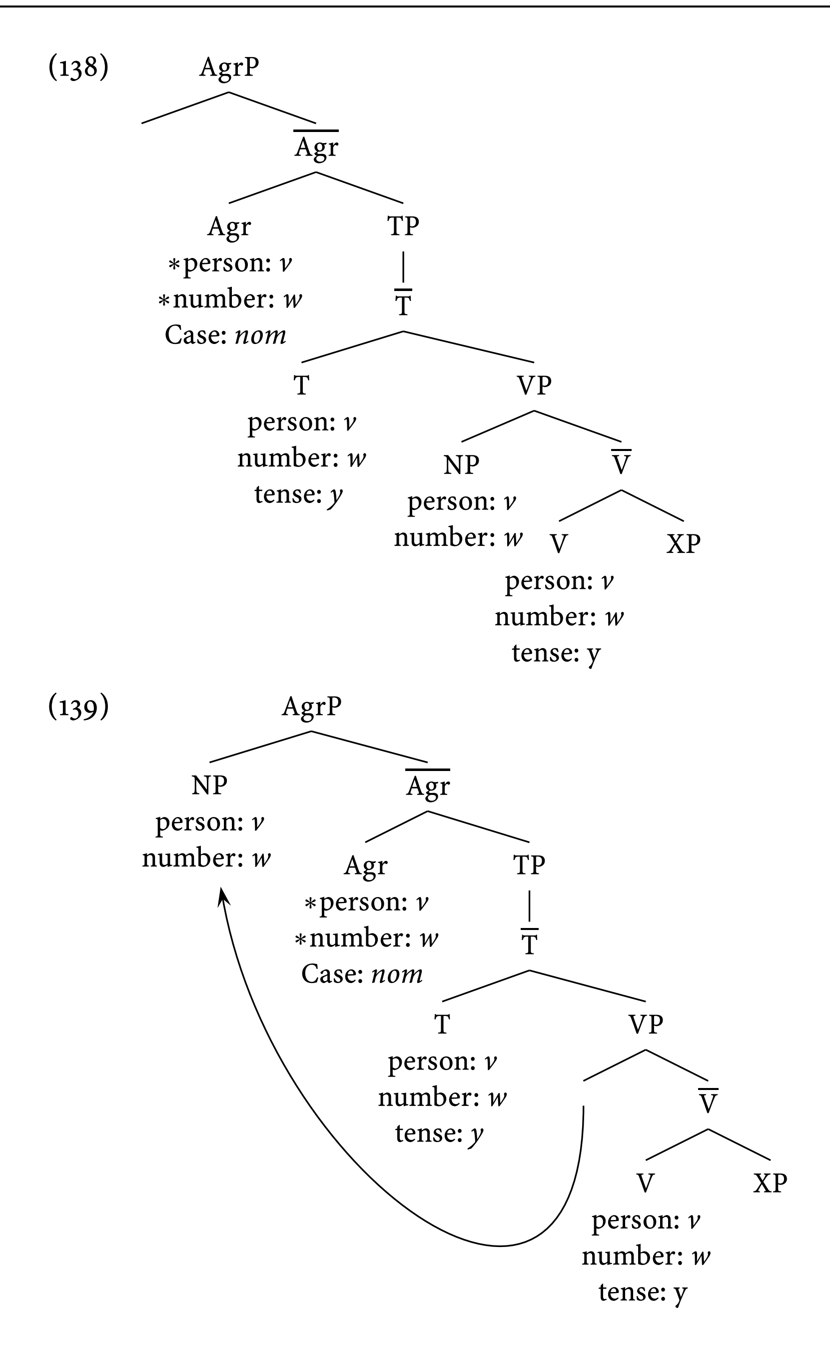
out, is agreement. We have already built into our system that verbs express the values of agreement features. But, as the label "agreement" hints at, those features are determined by the nominative Case marked NP. As the table in (137) indicates, the number and person of the nominative Case marked NP in English controls the morphological form of the finite verb.

(137) For present tense *be*

	singular	plural
1st person	*I am*	*we are*
2nd person	*you are*	*you are*
3rd person	*he/she/it is*	*they are*

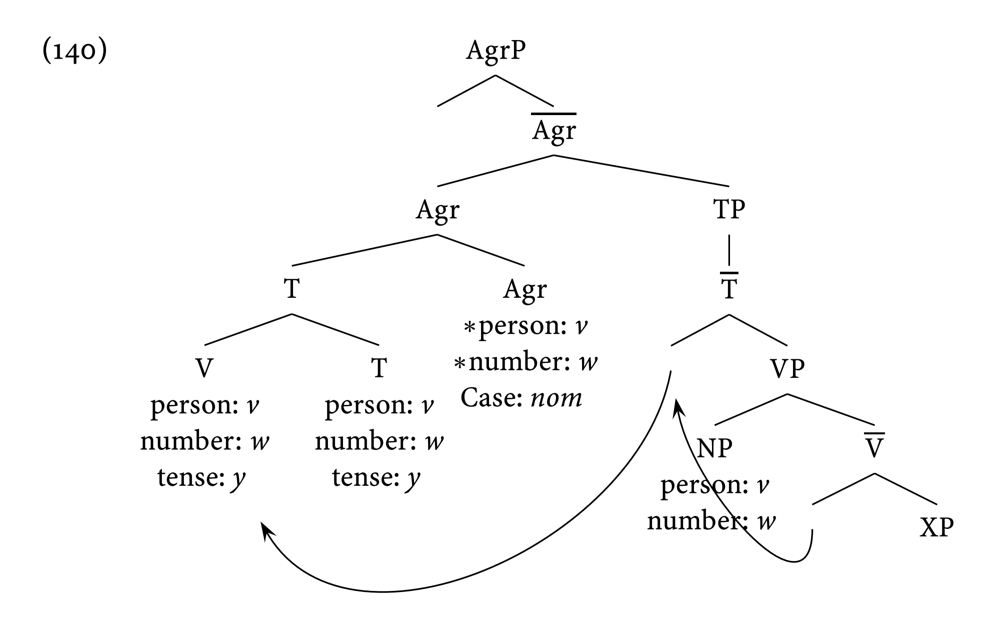
Indeed, all other verbs merely distinguish the 3rd singular forms from all others. But the existence of the paradigm in (137) indicates that both person and number features are present in the agreement morphology of English. We should therefore decompose the agreement feature into number and person features and let AGREE fix the values of these features in a way that involves not just Agr^0 and the verb, but the nominative Case-marked NP as well. We'll take up that task in the next section.

The proposal in Alexiadou and Anagnostopoulou (1998), then, requires that it be the agreement feature that has the ∗ property responsible for manufacturing EPP effects. Schematically, then, the relevant verb initial languages will have a D-structure representation like that given schematically in (138) below. In one class of language, the requirements of the ∗ attribute are satisfied by Argument moving the NP into Specifier of AgrP. This produces the typical word-order in English sentences, illustrated by (139).



In the other class of language, the verb's movement satisfies the ∗ property on the person and number features, resulting in an S-structure like (140). As can be seen, the correlative properties of high verb and low nominative NP are captured.

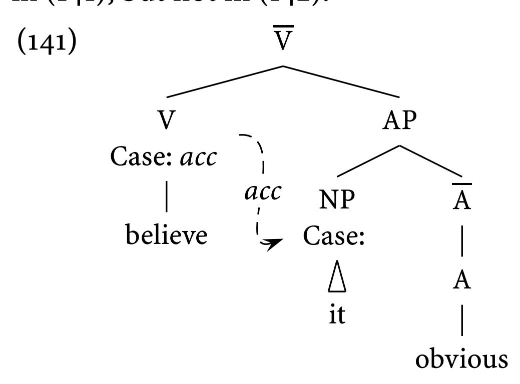
(140)



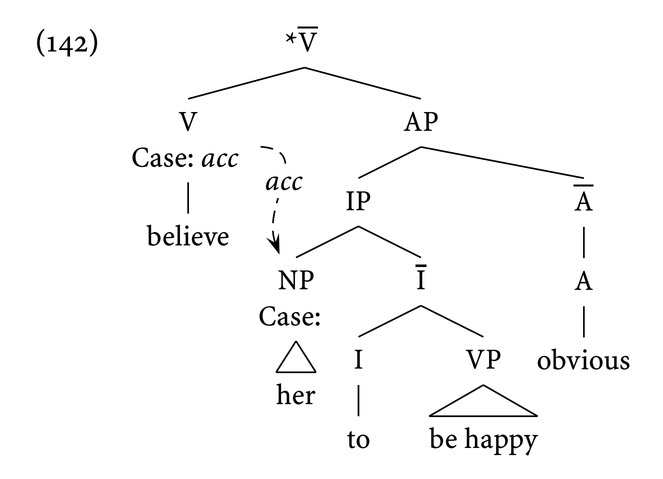
The cases in English involving expletive *there* seem to form a special case. Like the languages with the syntax in (140), the movement of the nominative Case marked NP is not required by the * property. But unlike the languages in which this is generally possible, the movement of verb does not satisfy the * property; instead, the presence of *there* does. We will not examine how *there* has this ability, nor will we try to unpack the causes for the constraints that govern this word-order in English. With these problems bracketed, however, we can see the *there*-construction in English as a special instance of the syntax that gives subject inversion elsewhere.

As noted above, this means that we will have to abandon our present formulation of klose, the locality condition on feature valuation when phrases are involved. We designed klose so that it would allow accusative Case assignment in (141), but not in (142).

(141)

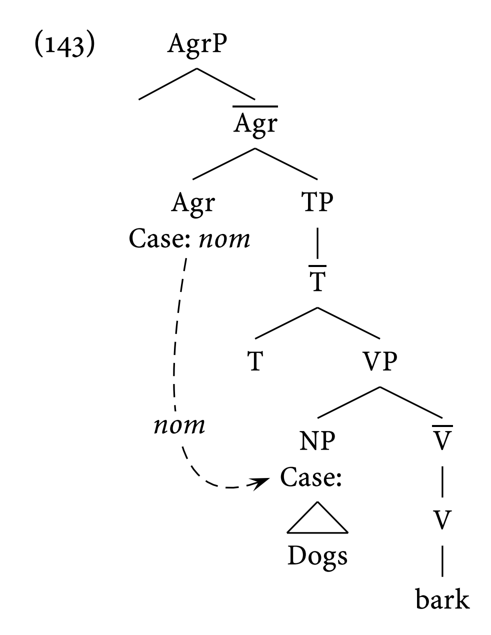


(142)



We now need it to allow (143) as well.

(143)



We cannot define klose in such a way that it legislates against too many phrases between the Case assigner and the NP that gets assigned Case, as we did originally.

One direction that some have taken is to make reference to the geometries of the phrase markers. In the good examples, both (143) and (141), the NP is embedded in phrases that are either complements or X^0s. In the bad example, by contrast, the NP is embedded within a phrase that stands in Specifier position. A wide array of movement, and other, relations seem to be blocked

when the reach into a phrase in a Specifier position. For instance, neither *wh* Movement nor Heavy NP Shift are possible out of such phrases:

(144) a. * Who have you said that a book about – arrived.
b. * That she eats – is obvious lukewarm natto rice.
compare:
She eats fervently lukewarm nattoo rice.

It appears that there may be something quite general about phrases in Specifier positions that makes them incapable of engaging in syntactic relations outside those phrases. Ross (1967), who charted out many such cases, calls these regions "islands."[15] One hypothesis about the ungrammaticality of (142), then, is that this has nothing to do with proprietary constraints on feature assignment but is the result of a global constraint on establishing relations into Specifiers. Let us adopt that idea. This means that we have not found any locality condition on the assignment of features to phrases. We revise AGREE to (145), then.

(145) Let X and Y be terms that carry features, and let the features on X have a value and those on Y be unvalued. The features on Y can get the value of the matching features on X just in case Y AGREES with X. X and Y AGREE iff:
a. X c-commands Y, and
b. If Y is a head, there is no Z^0 such that X c-commands Z^0 and Z^0 c-commands Y^0.

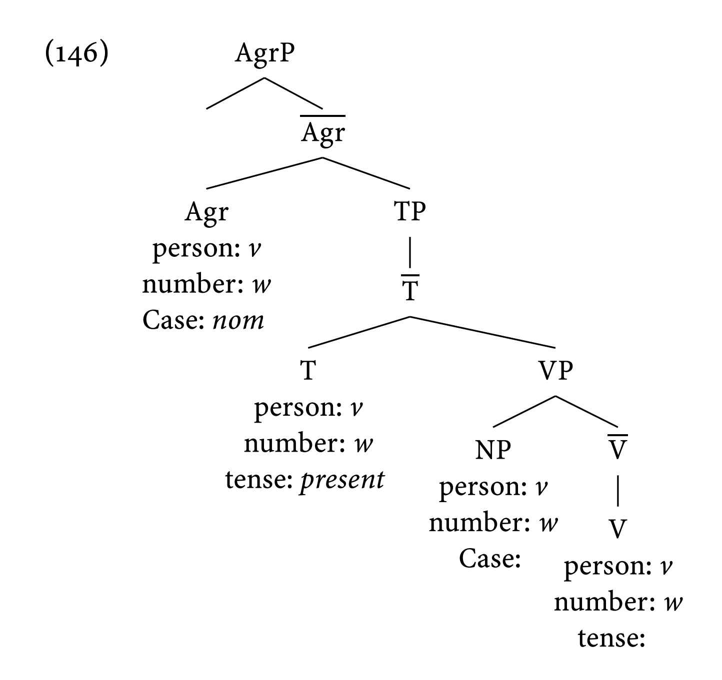
Let's now turn to unpack agreement into the person and number features we've seen them consist in, and work out how AGREE can match the values of these features on Agr^0, a verb and an NP.

4.6.2 Case and Agreement

Recognizing that the person and number features make up agreement, and that they are given the same values on both the finite verb and the nominative Case-marked NP, we want to distribute these features as indicated in (146) below.

15 The particular island involved here goes by the names: "Subject Condition" and "Generalized Left Branch Condition." For two popular approaches to defining this island, see Huang (1982), Lasnik and Saito (1984) and Chomsky (1986a), for one, and Kayne (1983) and Pesetsky (1982) for the other.

(146)

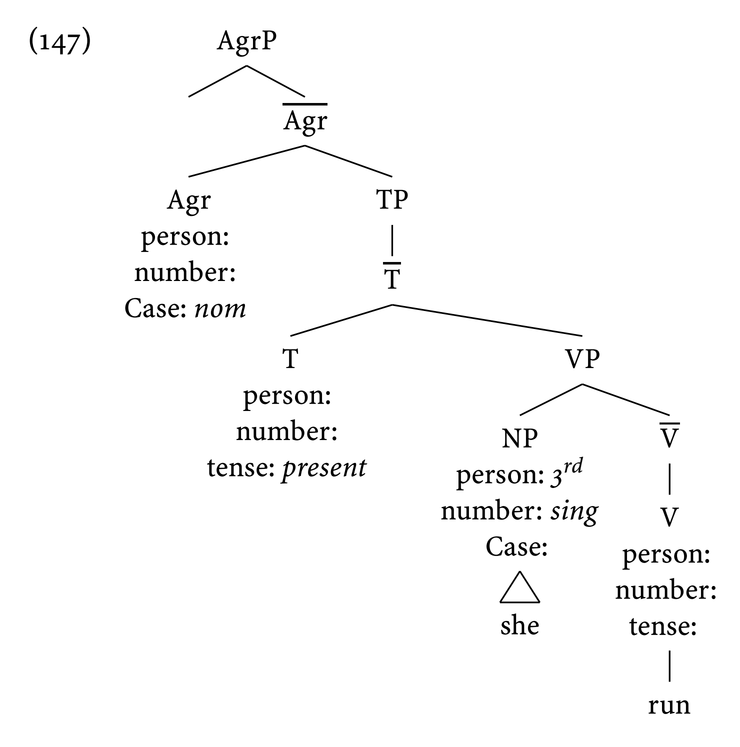


Let's consider now how AGREE would cause the values for the person feature (*v* in (146)) and the number feature (*w* in (146)) to be fixed.

It seems natural to think of these features as coming valued on the NP involved in the relation. These features have a role in the semantic interpretation of the NPs, though they don't obviously have any semantic role on Agr^0 or the verb. For this reason, we should take the D-structure representation to be as (147) below, which maps onto the sentence *she runs*, indicates. Under our present definition of AGREE, values are passed from one item, α, to another, β, only if α c-commands β. This means that the person and number values on NP cannot be assigned to Agr^0 in (147), though it can assign those features directly to the verb. This definition of AGREE, then, is not compatible with there being person and number features on Agr^0 if those features come valued on NPs. But if we do not put person and number features on Agr^0, then we will no longer be able to give these features the ∗ attribute and employ the account for the subject inversion typology that we reviewed in the previous section.

In fact, there is another reason to think that Agr^0 has person and number features associated with it. There is a close connection between the features that figure in agreement — in English: person and number — and those that are responsible for Case. Features that can enter into agreement relations are called "ϕ features," and in addition to person and number, they also include gender and, perhaps, word class markers like those that play a role in Bantu

(147)



languages and elsewhere. It is very common cross-linguistically that the NPs which bear nominative Case also agree in ϕ features with the verb. Moreover, the distribution of agreement and nominative Case within, and across, languages is very similar. Within English we see this in the fact that those NPs which bear (overt) nominative Case do so in the very clauses where they trigger (overt) agreement. Nominative Case marked NPs are exclusively found as "subjects" of finite clauses, and this is where agreement triggering NPs are found as well. This is extremely common: in those languages that have both nominative Case assignment and agreement, it is the nominative Case marked NP that enters into agreement. This correlation is also reflected in those languages that have only nominative Case assignment or agreement. Just as in English, the nominative Case, or agreement, targets certain NPs in finite clauses but not in other reduced clauses.

This relationship is found with respect to accusative Case as well. There is no agreement associated with accusative Case-marked NPs in English, but in many languages there is. This so-called "object agreement" has a distribution that looks very much like that which accusative Case-marked NPs have in English: it depends on the choice of verb ("transitive" verbs trigger it and "intransitive" verbs don't) and it targets NPs that are roughly in the positions that

accusative Case marked NPs are in English.

What these correlations suggest is that there Case and agreement are always controlled by the same term, and that when both Case and agreement are involved, they are bundled together by that one term. I'll describe these correlations with (148).[16]

(148) If a language has an unvalued Case feature or unvalued ϕ features, then there will be some head in which they are all bundled.

ϕ features might well exist in languages where there is no agreement phenomenon, in which case they would only play a role in determining the meanings of the NPs they are in. In such languages, all the ϕ features would be valued. The first clause in (148) therefore restricts the constraint to those languages that have Case marking or agreement. It requires in such languages that there be a head which mediates the values of those features, thereby gluing them together. In English, this head is Agr^0.

For these reasons, then, we should leave person and number features associated with Agr^0. As noted above, this means that unless we decide against letting the valued versions of the person and number features come with the NP, we have a problem for our formulation of AGREE. We cannot restrict AGREE so that it only passes values down to c-commanded terms. In (147), repeated below, we need to let the values for person and number on the NP be passed up to Agr^0. To this end, AGREE becomes (149).

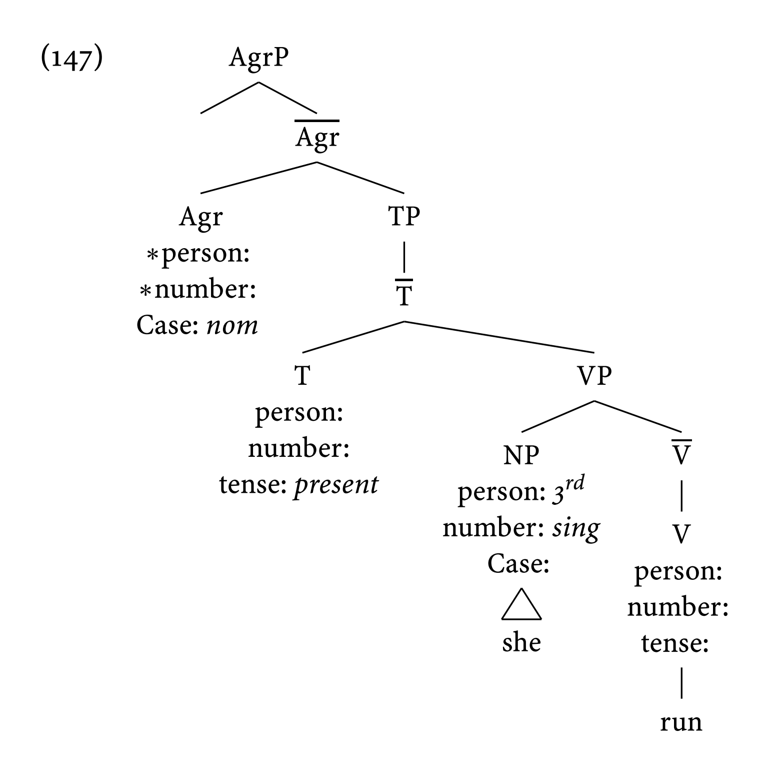
(149) Let X and Y be terms that carry features, and let the features on X have a value and those on Y be unvalued. The features on X or Y can get the value of the matching features on the other just in case X and Y AGREE. X and Y AGREE iff:

a. one of X or Y c-commands the other, and
b. If Y is a head, there is no Z^0 such that X c-commands Z^0 and Z^0 c-commands Y^0.

This version of AGREE still requires that the two terms in an AGREE relation also be in a c-command relation. But it doesn't require that the values be passed from the c-commander to the c-commandee. This will have the desired effect,

16 This claims that if a language has both Case assignment and agreement, only one of the Case assigners will be forced to trigger agreement with the NP it assigns Case to. What we (probably) want is to force every Case assigner to control agreement with the NP it assigns Case to. We want to allow languages like English, in which one Case assigner (nominative) is tied to agreement but another (accusative) isn't, while still tying together the Case assignment and agreement phenomena.

(147)

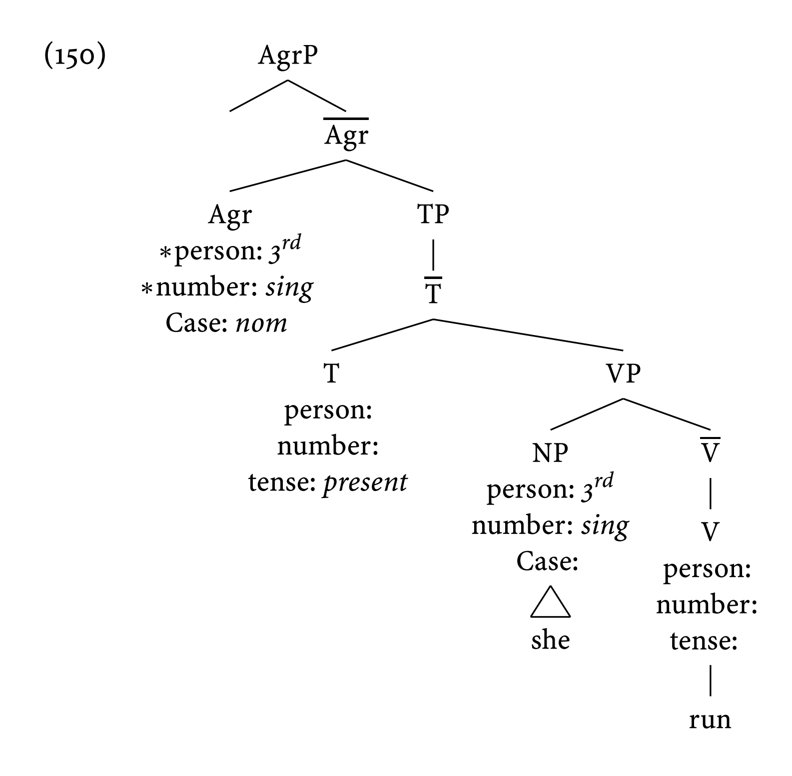


then, of letting the values for person and number features on an NP be assigned to an Agr^0 that this NP does not c-command. I will allow AGREE to convert (147) to (150) below.

The version of AGREE in (149) would also let the values that NP has for the person and number features be fixed on T^0 and V^0 by putting these terms into an AGREE relation with NP. This isn't desired, however. If the NP could directly agree with verbs, and other terms, then our desire to tie agreement to Case would be lost. We would expect that the values for ϕ features on a verb could be valued by an NP that is not getting its Case from Agr^0. What we see is something different, however. It is the very NP that bears nominative Case whose ϕ features arise on the verb. We must prevent the NP from entering into an AGREE relation with V^0. Its values for person and number must come from Agr^0. A common way of ensuring this is to require that the c-commanding term in an AGREE relation is a head. This would prevent the NP in (149) from entering into an AGREE relation with the verb that follows. We therefore change AGREE once more.

(151) If X c-commands Y, then the unvalued features on X or Y can get the value of the matching features on the other just in case X and Y AGREE.

(150)



X and Y AGREE iff:

a. X is a head, and
b. If Y is a head, there is no Z^0 such that X c-commands Z^0 and Z^0 c-commands Y^0.

The c-commanding head that enters into an AGREE relation is often called the "probe," and the other term in the AGREE relation is known as the "goal." What (151) insists, then, is that the probe be a head.

4.6.3 A constraint on Agreement

Now that the definition of AGREE has matured to the point where it applies correctly to the basic cases of agreement and Case assignment, we can discover a locality condition on AGREE when the goal is a phrase. We have already discovered a locality condition on AGREE when the goal is a head. In that scenario, recall, we see the effects of the "Head Movement Constraint": no other head may come between the probe and goal. We will now see a reason for believing that something similar arises when the goal is a phrase.

Imagine a D-structure representation like (152), where the imaginary verb, *blort*, is transitive.

(152)

```
AgrP
├── (Specifier, empty)
└── Agr-bar
    ├── Agr
    │     *person:
    │     *number:
    │     Case: nom
    └── TP
        └── T-bar
            ├── T
            │     person:
            │     number:
            │     tense: past
            └── VP
                ├── NP
                │     person: 3rd
                │     number: sing
                │     Case:
                │     △ she
                └── V-bar
                    ├── V
                    │     person:
                    │     number:
                    │     tense: past
                    │     | blort
                    └── NP
                          person: 3rd
                          number: sing
                          Case:
                          △ it
```

The verb *blort* is one of the kinds that Burzio's generalization describes as not occurring: it is a verb that c-selects two NPs but does not come equipped with a valued Case feature. We saw in the previous chapter that a way of understanding why such verbs should not arise is by considering how the D-structures they produce would satisfy the Case filter. Under a certain set of assumptions about how Case is assigned, we discovered that such D-structures could not satisfy the Case filter. With no well-formed S-structure able to express these verbs, their non-existence is explained.

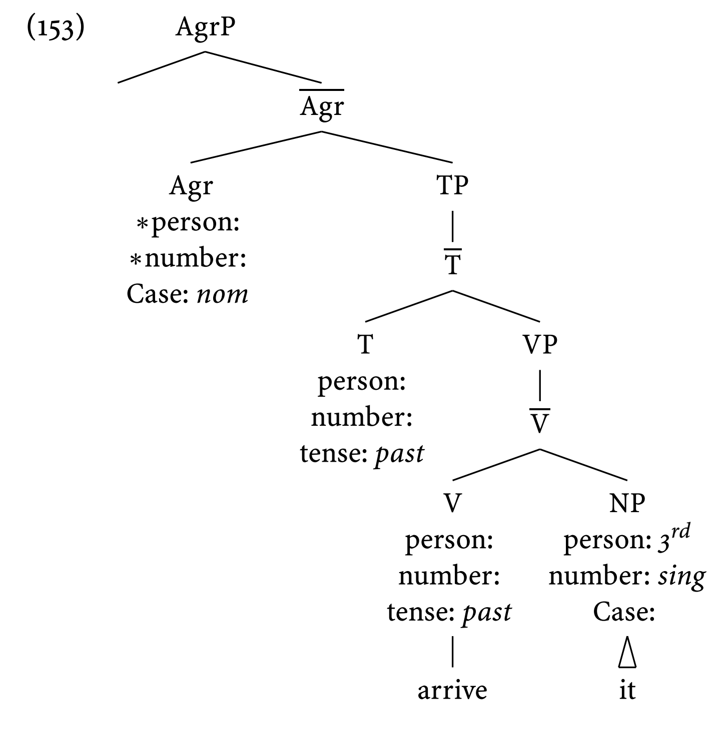
One of the assumptions necessary to block (152) is that the NP in object position is not capable of getting nominative Case from Agr^0. At the time we examined this scenario, we were working with a system of Case assignment that required NPs to move into Specifier of AgrP to get their Case feature valued. Now that this system has been revised, we should reconsider how (152) is prevented from having a good S-structure associated with it.

The definition of AGREE in (151) does not have this desirable consequence. The unvalued Case features on both NPs could get a value from Agr^0. To prevent this, we have two options. We could prevent the Case feature from entering

into more than one AGREE relation, thereby preventing it from giving its value to both of the NPs in (152). This would force a violation of the Feature Valuation Requirement, the descendant of the Case filter, for one of these NPs. The other option is to place a locality condition on the AGREE relation that would prevent Agr^0 from passing its Case feature value to the object NP. (A third option is to invoke both options.) Both of these ideas have been pursued. I believe it is easiest to find independent support for the second option, so I will sketch how it could be formulated.

We want D-structure representations like (153) to produce well-formed S-structures, so we cannot design a system that would always prevent an object from getting its Case feature valued by Agr^0.

(153)

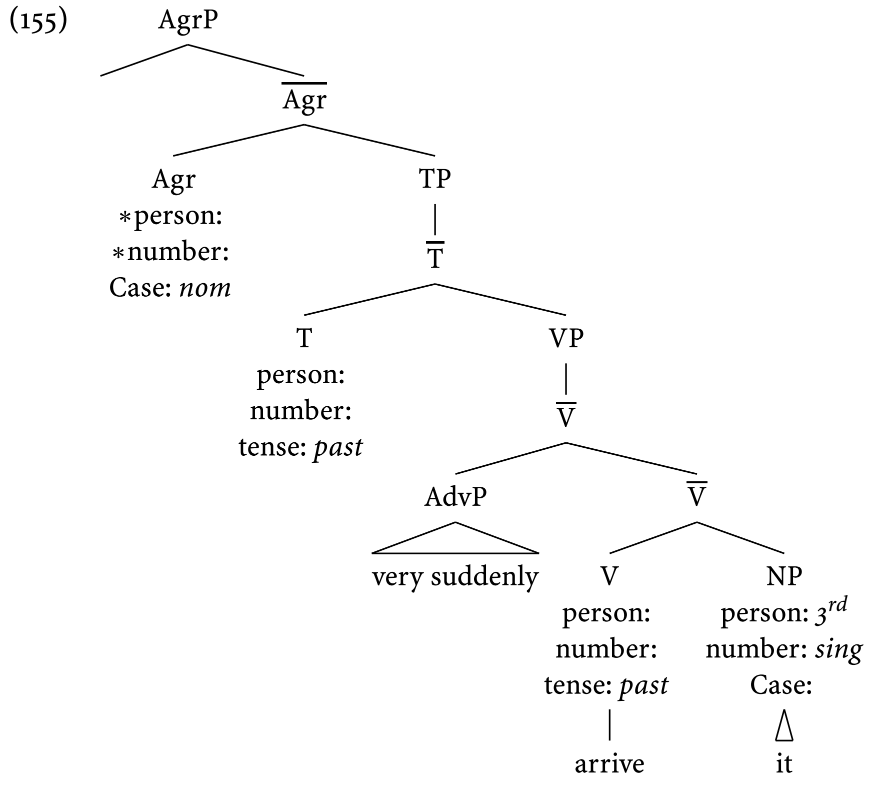


We want to make the presence of the NP in Specifier of VP responsible for blocking Case assignment to the object NP in (152). One way of doing this is with (154).

(154) If X c-commands Y, then the unvalued features on X or Y can get the value of the matching features on the other just in case X and Y AGREE. X and Y AGREE iff:

a. X is a head, and

(155)



b. There is no Z such that X c-commands Z and Z c-commands Y, where Z and Y match in projection level.

This definition has generalized the Head Movement Constraint clause in AGREE to phrases. It prevents there from being any term between the probe and goal that is of the same projection level as the goal. If the goal is an XP, then there can be no closer NP to the probe; and if the goal is an X^0, there can be no closer X^0 to the probe. This will correctly prevent (152) from having a grammatical outcome as it will prevent the object NP from getting a Case value from Agr^0 since there is a closer NP, the subject.

The locality condition in (154b) is too strong. It prevents any phrase from standing between an NP and the head it gets its Case valued from, and that would prevent (155) below from having a grammatical outcome. But (155), of course, leads to the well-formed S-structure: *It very suddenly arrived.* Let's weaken the locality condition, then, to (156).

(156) If X c-commands Y, then the unvalued features on X or Y can get the value of the matching features on the other just in case X and Y AGREE. X and Y AGREE iff:

a. X is a head, and

b. There is no Z such that X c-commands Z and Z c-commands Y, where Z and Y the same kind.

c. X and Y are the same kind if they are both heads, or they are both NPs

It is probably possible to come up with a more natural definition of kind, but this will suffice for our purposes.

This way of bringing the locality condition that arises when the goal is a head with the locality condition that arises when the goal is a phrase comes from Rizzi (1990). He framed his condition as a constraint on movement since the cases he examined were all ones in which movement was triggered by the AGREE relation. He named his condition "Relativized Minimality." Translated into a condition on AGREE, it is often called "closest attract."

4.7 Summary

Here is a brief recap of the elements we have introduced to our grammar in this chapter.

(157) a. If a feature, α, has $*$, then it must be pronounced with the term whose features it is in an AGREE relation with adjoined to a projection of the head containing α.

b. If a language has an unvalued Case feature or unvalued ϕ features, then there will be some head in which they are all bundled.

c. If X c-commands Y, then the unvalued features on X or Y can get the value of the matching features on the other just in case X and Y AGREE. X and Y AGREE iff:

i. X is a head, and

ii. There is no Z such that X c-commands Z and Z c-commands Y, where Z and Y the same kind.

iii. X and Y are the same kind if they are both heads, or they are both NPs

d. Feature Valuation Requirement (holds of S-structure)

The values of a feature must be morphologically expressed and every feature must get a value.

I've changed the wording of some of these conditions in order to fit them together better. For instance, the formulation of the $*$ property that I set out

had the ∗ property of heads, rather than of features. But in order to embrace the account for subject inversion in Alexiadou and Anagnostopoulou (1998), we found it necessary to see the ∗ attribute as property of features. There are other, similar, small rewordings in (157).

5

Determiner Phrases and Noun Movement

One of the puzzles we stumbled over in introducing phrase structure rules involved the internal shape of determiner phrases. I noted that the set of strings that constitute DPs is miserably anemic. There are very few examples of non-trivial strings of words that offer themselves as possible determiner phrases. Typically, a determiner phrase appears to be constituted of nothing more than its head. The sorts of examples I offered as candidates for this family of strings were things like the bold-faced material in (1).

(1) a. **all but three** determiners
    b. **more than six** children
    c. **two dozen** eggs

But there are reasons to think that these examples don't have parallel parses, and that, in fact, none of them fit to a DP string in the desired way. It's probable that *dozen*, in (1c) is an adjective; this can be seen by observing that it can follow other adjectives (something determiners aren't capable of doing):

(2) an expensive dozen eggs

(1a) involves a coordinator, *but*, which will invoke the sorts of structures we have encountered before with coordinations. (1b) involves what is known as a "comparative construction," whose syntax, like that with coordinations, invokes larger joined structures. We won't examine these cases in any detail here, but let me offer as a way of thinking about the syntax of these cases that makes their semantics transparent, something along the lines of (3), where the

~~struck-out~~ material should be understood as syntactically present, but phonetically absent.[1]

(3) a. [all ~~of the determiners~~] but [three determiners]
 b. [more ~~of the children~~] than [six children]

Imagine, that is, that these cases involve bringing two full NPs together, and that a process of ellipsis removes the $\overline{\text{N}}$ from the first NP and, moreover, this $\overline{\text{N}}$ is understood to refer to the same set of individuals that the $\overline{\text{N}}$ in the other NP refers to.

If these cases don't involve strings that have the same distribution as determiners, then where are the strings that are determiner phrases? Why are they so hard to find?

This problem can be related to another, worse, problem. Remember that determiner phrases compete with genitive NPs for the Specifier of NP position. This is what is responsible for the paradigm in (4).

(4) a. Mary's lamp
 b. the lamp
 c. * the Mary's lamp

We adopted a view of NPs that embraced a constraint — yet to be found — that limited DPs and genitive NPs to their Specifier position. When we transited from a model about syntactic form that used Phrase Structure rules to one that involved $\overline{\text{X}}$ Theory and other general principles, we lost a way of expressing this fact. We must now understand how to capture the competition between determiners and genitives for the "first" position in NPs. This involves understanding what it is that governs the distribution of genitive NPs and determiners.

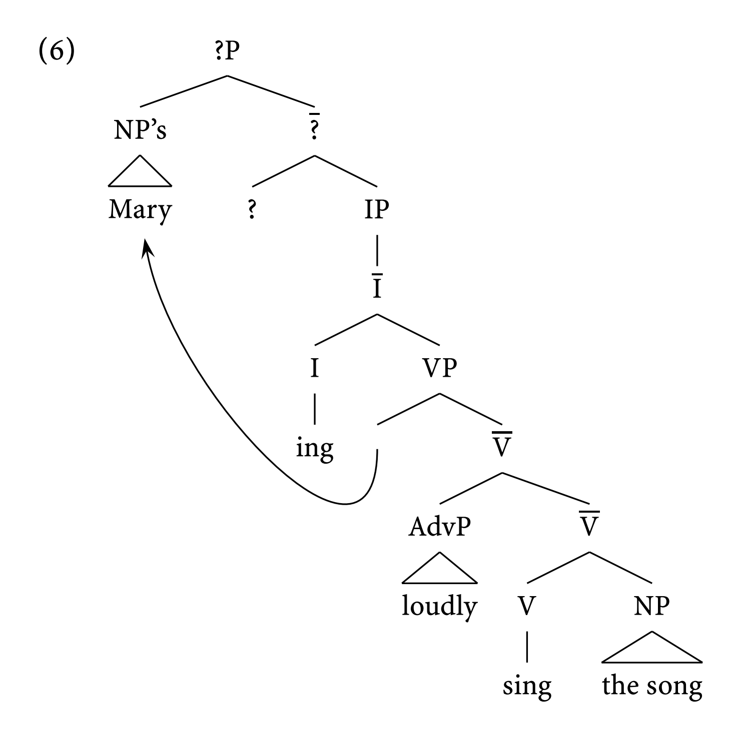
A problem for capturing the related distributions of determiners and genitives arises in certain cases which look rather like clauses, but which nonetheless have genitive subjects. These are called "gerunds," and (5) provides some examples.

(5) a. [Mary's loudly singing the song] bothered us.
 b. I recalled [Mary's fixing the car].
 c. [Mary's having talked to John] wasn't widely known.

This suggests that these phrases have an organization something like (6).

1 For an examination of cases like (3b), see Hackl (2000).

(6)



But what's the "?" in this graph?

5.1 The DP Hypothesis

There is some evidence that ?P has the same distribution as NPs. Recall that NPs are subject to the Case Filter, and as such, are unable to stand after adjectives, which apparently are incapable of assigning Case. The same is true for these sorts of expressions:

(7) a. I was happy with Mary's singing the song.
b. * I was happy Mary's singing the song.
(compare: "I was happy that Mary sang the song.")

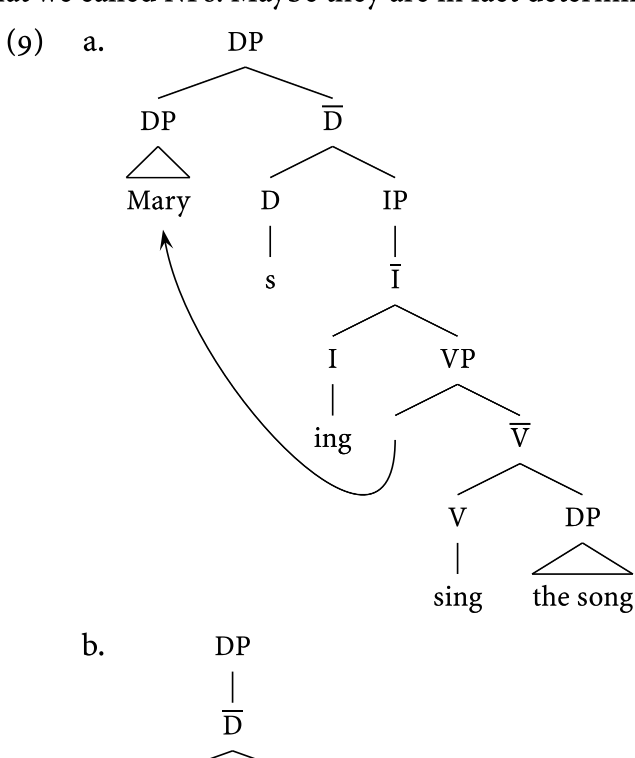
And these expressions can be conjoined with NPs, which, if we've got the constraints on coördination correct, also indicates that they are NPs.

(8) [Mary's singing the song] and [my subsequent departure] enraged the organizers.
(compare: "*[Mary's singing the song] and [that I subsequently departed] enraged the organizers.")

But if ?P is a noun phrase, then the law of endocentricity demands that ? be a noun, contrary to fact. Something's amiss.

One way of thinking about this problem goes as follows. What makes the distribution of ?P look like that of noun phrases is the presence of the genitive NP. So, maybe we should call "?" the head that determines genitive Case on the NP which surfaces in its Specifier. Maybe, actually, it is the genitive *s* itself. This would mean that the distribution of Genitive Phrases is the same as NPs (perhaps). And since Genitive NPs are in complementary distribution with determiners, maybe we should rethink how we earlier characterized the phrases that we called NPs. Maybe they are in fact determiner phrases, as in (9).

(9) a.



b.

DP

D̄

D the

NP

N̄

N

lamp

If this is correct, it would also answer the problem we began with. The reason DPs look so anemic is because they're considerably larger than we thought.

The two considerations I've just adduced in favor of reanalyzing NPs as DPs with NPs inside them can be found in Abney (1987).[2] He gives another, smaller, argument on behalf of this reanalysis that relies on a mystery concerning the expression of Adjective Phrases in English. The mystery is that there is a constraint on Adjective Phrases in English nominals which determines how large they can be depending on whether they precede or follow the noun. As (10) shows, when an AP has nothing but its head in it, it prefers preceding the noun, whereas if it contains material following the A^0, it prefers following the head.

(10) a. some angry children
b. * some children angry

(11) a. * some angry at Bill children
b. some children angry at Bill

There is a systematic exception to this, and these are expressions like *everyone/everything*, *someone/something*, *anyone/anything* and *no one/nothing*.

(12) a. someone angry
b. something large
c. * angry someone
d. * large something

(13) a. everyone angry
b. everything large
c. * angry everyone
d. * large everything

(14) a. no one angry
b. nothing large
c. * angry no one
d. * large nothing

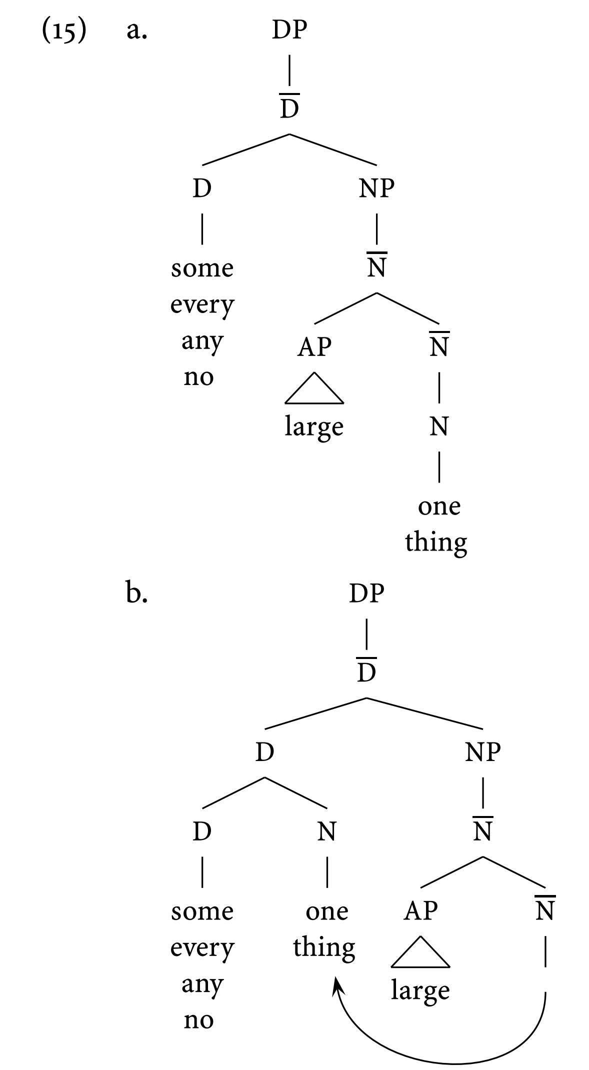
Abney[3] suggests that an analysis of this exception should not make it accidental that the determiners *every*, *some*, *any* and *no* and the nouns *one* and *thing*

2 I've changed slightly his discussion of cases like "Mary's singing the song" — but the spirit of the argument is his.

3 Who is here following a suggestion of Richard Kayne's, who in turn is building on ideas in Postal (1969).

are involved. More particularly, it should not be accidental that the only expressions in English which seem to be made up of a determiner and noun sequence should be the very expressions which seem to violate this generalization. He recommends that we see these cases as coming about through movement of *one/thing* onto the determiner; that is, he suggests that (12) be involved in a derivation that includes the parses in (15).

(15) a.



Now Head Movement can combine one with D^0 to form the DPs in (12).

Further, to the extent that combining *one* with *some/every* really involves Head Movement, we have an argument for the reanalysis of NPs into DPs. This is because the Upwards Constraint and the Likes Attracts Likes constraint com-

bine to allow *one* to adjoin only to a head that c-commands its original position. So, if the *some* and *every* parts of *someone* and *everyone* are determiners, and the *one* part is a noun that has Head Moved into these determiners, then it must be that D^0 c-commands NP.

If these considerations are on the right track, it demands that we change our way of talking about nominals altogether. Everything we once thought to be true of NPs, is now true of DPs instead. So, for instance, the Case Filter is something that fixes the position of DPs, not NPs. NPs are now found pretty much only inside DPs and not, as we previously thought, in subject and object positions. From this point forwards, then, let everything that we have credited to NPs hold of DPs instead, and let NPs be selected only by determiners, thereby fixing their position within DPs.[4]

We have also seen, faintly, evidence that nouns move internal to DPs in a way somewhat like the movement of verbs internal to CPs. Indeed, there is a variety of interesting evidence that Noun Movement exists to a larger degree than just that found in the *someone* and *everyone* cases mentioned above. Moreover, there is some evidence that this movement relocates a noun to a head associated with inflectional morphology, much like the situations we have viewed involving verbs and inflectional morphology. In gross terms, then, DPs and IPs have a variety of parallels; it is this parallelism that Abney focuses on in the first two chapters of his dissertation.[5]

In this chapter, we will examine a paradigm of facts which focus on the noun movement part of this parallelism.

5.2 Noun Movement

One paradigm of facts that has been accounted for in terms of noun movement concerns a difference in the position of a noun's "Subject," which is how we might characterize the terms that appear as genitives in English. In Romance (by which I will mean here Catalan, standard French and Italian), the "subject" argument can appear between the noun and its complements.

4 We will quickly see that there is evidence that there are phrases that DPs embed that in turn embed NPs, and, consequently, there is reason to believe that D^0 does not select NP. If this evidence is correct, we'll need instead to let D^0s select the phrase whose head selects NP. In general, what we'll want is to guarantee that the terms which select NPs are always found within DPs (unless, of course, we discover that NPs can be found elsewhere).

5 Abney (1987).

(16) a. L'opinione di Maria di Gianni
the opinion of Mary of John
'Mary's opinion of John'
(Italian)

b. les novel.les d'en Pere de Maria
the novel of Pere of Maria
'Pere's novel of Mary'
(Catalan)

c. le portrait de chaque peintre étranger de son enfant
the portrait of each painter foreign of his child
'the picture by each foreign painter of his child'
(French)

This could be made sense of if we adopt the Derived Subjects Hypothesis for nominals as well as for clauses, and suppose that there is N^0 movement in Romance but not English. If we understand the Derived Subjects Hypothesis to claim that it is the highest $\overline{N}$ which assigns the θ-role that "subjects" in DPs receive, then this will put these subjects in Specifier of NP underlyingly. If nouns then move leftwards in Romance, and the subjects of DPs can remain in their underlying position, this will have the desired consequence of placing nouns to the left of the subjects.

I think the first argument of this sort comes from Cinque,[6] who makes the argument based on the position of "ethnic" adjectives, which also can be found postnominally in Romance.

(17) L'invasione tedesca dell'Austria.
the invasion german of Austria

Ethnic adjectives seem able to bear a subject θ-role assigned by a noun. So, consider, for example, the contrast in (18) below.

(18) a. the American car in the showroom
b. the American waltz on the radio
c. the American opinion of the blockade
d. the American discussion of trade barriers

6 In an unpublished talk delivered at the Università di Venezia in 1990. See his GLOW abstract from the 1992 Lisbon meeting.

In (18a), *American* serves as a genuine modifier, merely attributing to the referent involved that it has the property of being American. But in (18b), *American* refers to an abstract entity that is constituted of the American people or the American government.[7] That this is a function of these adjectives bearing the external θ-role which *opinion* and *discussion* assign is indicated by the fact that this meaning is lost when there is another external θ-role bearer in the nominal.

(19) a. Uganda's American opinion of the blockade
b. Morocco's American discussion of trade barriers

Like other external θ-role bearers, then, this one shows up postnominally in Italian (and the other Romance languages), which can be accounted for if we let these adjectives be underlying placed in Specifier of NP, and then make nouns move leftwards past them.

There is an alternative method of generating these word orders. Giorgi and Longobardi (1991) suggest that there is a difference in the direction that Specifiers can branch in Romance and Germanic, and that this is responsible for the fact that the phrases that appear in these Specifiers, i.e., "subjects," show up following the noun in Romance but not Germanic. This alternative account predicts that postnominal "subjects" can follow the complements, and this is generally possible too. Thus the *di/de* phrases in (16) are actually ambiguous; either of them can have the subject or object reading. Giorgi and Longobardi suggest that this word-order alternation arises by virtue of a rule that moves the "object" past the right-branching subject position. The Noun Movement account would have to claim that the subject can move rightwards past the object.

There are reasons for doubting that the Giorgi and Longobardi account is correct, and this direction has largely been abandoned in the literature. One of these is that, as Valois (1991a) and Bernstein (1993) note, "ethnic" adjectives cannot follow complements in Romance.

(20) *L'invazione dell'Austria tedesca
the invasion of Austria german

(Valois 1991a, p. 374)

This can be related to the fact that ethnic adjectives seem unable to move. There is evidence in English for this which comes from the fact that ethnic

7 This quality of ethnic adjectives is first discussed, to my knowledge, in Kayne (1984, Chapter 7).

adjectives are unable to undergo the passive-like operation that nominals support in examples like (21), compare (22).

(21) a. Iran's bombardment by Russia took several weeks.
b. Uganda's invasion by Tanzania grinds slowly on.

(22) a. * The Iranian bombardment by Russia took weeks.
b. * The Ugandan invasion by Tanzania grinds slowly on.
(with an object interpretation for the adjective
(roughly Kayne's 1984, (32) and (33), p. 139)

As we shall have occasion to see, there is evidence that the genitives in (21) have undergone A Movement from a position to the right of the noun, where they receive their θ-role. Ethnic adjectives, apparently, are unable to move from this position. Instead, they are stuck in the position from which they get their θ-role. Thus, the fact that they appear in Romance between the noun and the noun's complements suggests that the underlying position to which the external θ-role is assigned in nominals is to the left of the complement. This follows if the Specifier of NP branches to the left rather than the right. This fact, then, fits the model of Romance nominals which has the noun moving leftwards past the subject.

So, now, where are the nouns moving in these cases? One possibility, explored in a variety of places, is that the intermediary position is where number morphology is associated. There is some prima facie reason for thinking that number morphology heads a syntactic phrase: Cross-linguistically this is common, as Dryer (1989) shows. Thus, in Yapese, for example, the plural/singular/dual categories are expressed with separate morphemes.

(23) a. ea rea kaarroo neey
sing car this
b. ea gal kaarroo neey
dual car this
c. ea pi kaarroo neey
plural car this

This at least suggests that Universal Grammar makes projecting a syntactic phrase above Number a possibility. Further, Dryer shows that the relative order of Num^0 and Noun correlates with Verb-Object word order. This would be explained, on standard theories of word order typology, if Num^0 is in a head complement relation with Nouns. Moreover, Dryer finds that most times there

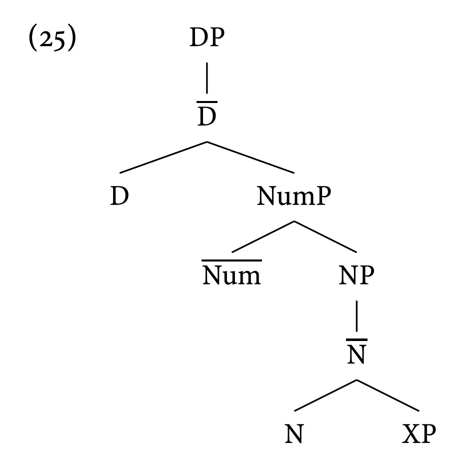
is a number word, it falls more embedded in the nominal than do determiners, but above adjectives and the noun. He provides examples like the following.

(24) a. ha ongo puha'e ua
art dual box two
(Tongan)

b. do mamu ragha
tree big plural
(Kimaghama)

c. me-ria rabiri
plur-new paddle
(Cayuvava)

There are exceptions, but this can be said to be the "basic" order among these terms. If so, the pattern that emerges can be sketched in (25) below. This is

(25)


what the statistical study yields.

Of course, if Dryer's conclusion that adjectives come between Num^0 and N^0 is valid for English, then English nouns must combine with this morphology in one of the ways we have seen possible in the verb/inflection cases. One possibility is that nouns overtly move to Number, but that this doesn't bring the noun to the left of the possessor in English because possessors are driven into Specifier of DP, an even higher position. Or, we might imagine that the noun undergoes covert movement to Number head.

So now what we want to determine is: Is there language internal evidence for the picture that Dryer gives us statistically? And, is there evidence that

bears on whether English differs from other languages in not enjoying overt N^0 movement to Number?

Let's tackle the second question first.

We have seen evidence for the movability of nouns. Is there language internal evidence that the site of this movement is Num^0. The best argument I know for this in the literature is found in Bernstein (1991), who manufactures what I called in the preceding chapter a "correlation argument." She claims that there is reason to believe that the position of nouns relative to adjectives correlates with the presence of number morphology on the noun. Her evidence comes chiefly from a comparison of Walloon and standard French. The contrast she describes is very like one that holds between English and French, however, so I will begin with an illustration of this difference.

In French, but not English (with the exception we've already noted), it is possible for single adjectives to follow the noun they modify.

(26) a. dès bêtes malades (French)
b. some sick animals (English)
*some animals sick

It is also possible to find prenominal single adjectives in French, as in the following example.

(27) a. une large vallée
a large valley
b. une vallée large
a valley large

But here Bernstein notes that there is a difference in meaning: in (27a), the nominal refers to an individual drawn from the set of things that are large valleys. In (27b), by contrast, a "member of a class of valleys which happens to be large" is denoted. In Giorgi and Longobardi's study of this phenomenon in Italian, they suggest that the prenominal depictive adjective can only get an appositive interpretation, whereas the postnominal one can have either an appositive or restrictive reading. The difference between an "appositive" and a "restrictive" reading is subtle. Roughly speaking, appositive modifiers contribute their meaning to the expression they are attached to in a way that is reminiscent of conjunction. So, for instance, in (28a) the PP *from Duluth* stands in the same relation to *Mary* as it does in (28b).

(28) a. Mary, from Duluth, has arrived.

b. Mary has arrived and she is from Duluth.

In an example such as (29), by contrast, *from Duluth* plays a more direct role in determining the reference of the DP it is attached to.

(29) Jill likes women from Duluth.

In this case, *from Duluth* restricts the reference of women to just those that have an attribute that Jill values: being from Duluth. One could not capture the meaning conveyed by (29) with a circumlocution, parallel to (28b), like:

(30) Jill likes women, and they are from Duluth.

Perhaps it is this sort of difference in meaning that correlates with the pre-nominal/post-nominal position of adjectives.

If so, it doesn't appear to always be the case, however. There are some examples where the alternation between Adj+N and N+Adj order doesn't appear to invoke any meaning difference. Valois (1991b) provides some examples in nominals with a deverbal noun.

(31) a. La probable invasion de Jupiter
the probable invasion of Jupiter

La invasion probable de Jupiter
the invasion probable of Jupiter

b. La fréquente invasion de Jupiter
the frequent invasion of Jupiter

La invasion fréquente de Jupiter
the invasion frequent of Jupiter

(Valois 1991b, 374)

Valois claims that there is no difference in meaning attendant with these word orders. What's going on here will have to await a better understanding of the syntax-to-semantics mapping of modification.

What is the source of the difference between French and English with respect to placement of these single adjectives. Why can they appear after the noun in French but not in English?

One possibility would be to blame whatever it is that prohibits bare adjectives from being right-adjoined to the nominal projection as the cause. Maybe this constraint is not present in French? Actually, however, there is evidence that this constraint is also present in French. We've seen that bare adjectives

can follow the noun, but they cannot follow the noun's complement, as in the following example.

(32) *Linvasion de Jupiter compléte

If bare adjectives could right-adjoin to a projection of a noun, there would be no reason for this asymmetry — they should be able to follow everything that is in an NP. But, on the other hand, if we assume that this constraint operates in French just in the same way that it does in English, then (32) will be ungrammatical for the same reason that the English version of this DP is. And, the fact that bare adjectives can follow single nouns, on the other hand, can be explained if we allow nouns to Head Move past adjectives in French, but not English.

Now, interestingly, Bernstein shows that Walloon patterns with English, and not with the other Romance languages, with regard to adjective placement.

(33) a. dés malâtès bièsses
the sick beasts

b. * dés bièsses malâtès
the beasts sick

(Walloon)

She suggests that the difference between French and Walloon is that the noun moves past the adjective in French but not in English or Walloon. And, further, she argues that the contrast between Walloon and French correlates with a difference in the way that number morphology is expressed on the noun. This, she argues, suggests that the position the nouns are moving to in French is a position associated with number morphology.

Her argument that number morphology is centrally implicated rests on several observations. First, she points out that neither colloquial French nor Walloon show (phonetically) the plural morphology that is orthographically present on nouns. Thus the bracketed portions of the following nouns are not pronounced.

(34) a. dès r'tchâfés crompîre[s]
the reheated potatoes

(Walloon)

b. des petites fille[s]
the small girls

(French)

These plural affixes are not present phonetically even when the context for liaison is provided.

However, she notes that in French there is a certain class of suppletive forms where the nouns do show a morphological mark for plurality. Examples are given below.

(35) a. un mal
an evil

b. des maux
evil(s)

(36) a. un oeil rouge
a red eye

b. des yeux rouges
red eye(s)

(37) a. quel cheval
which horse

b. quels chevaux
which horses

In Walloon, however, these forms always appear just in their singular form.

(38) a. on mâ
evil

b. dès mâ[s]
evils

(39) a. on rothch[e]-oûy
red eye

b. dès rodj[e]-z-oûy
red eyes

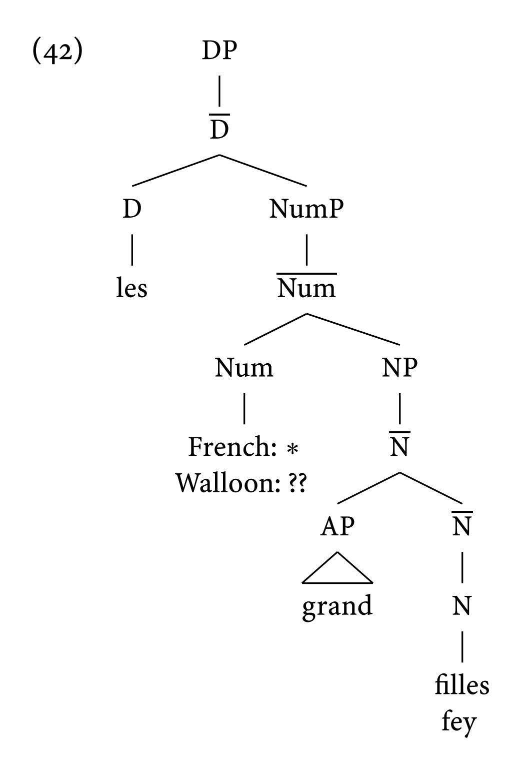
(40) a. [keː dzvaː]
which horse

b. [keː dzvaː]
which horses

Further, she notes that liaison is possible in literary French, but never in Walloon.[8] Thus, the orthographically present, but normally not phonetically manifest, plural suffix is pronounced in literary French in contexts like the following.

(41) Les train-z-ont du retard.
the train-s-are running late

From these data, she concludes that the plural affix in French is *-es*, or a suppletive trigger, and that it is absent altogether in Walloon (well, almost — we'll revise this immediately). Thus, using the inflectional feature model we adapted from Chomsky in connection with verbal syntax, this gives to French/Walloon nominals the D-structure in (42) below.

(42)


In French, the head noun is driven overtly into the Num^0 position to satisfy the ∗ attribute associated with the feature residing there, and this will bring it

8 But it's not that liaison is completely absent in Walloon — Bernstein notes that it is still present after plural determiners and pronouns (for this, see Bernstein (1991, note 7, p. 107).

past the adjective. This movement is blocked in Walloon because either there is no NumP, or its head is associated with a feature that has no ∗ attribute. Thus the availability of plural forms in French is correlated with its ability to appear before single adjectives. This achieves the desired correlation between presence of number morphology and N+Adj word order, and also supports the idea that number morphology is associated with an inflectional category that projects its own phrase.

Is it possible to tell whether Walloon has a NumP, or whether it is absent altogether? Bernstein suggests that there are reasons for assuming that Walloon does have NumP and, moreover, there is some reason to believe that it is actually filled with morphology. If this is correct, the crucial difference between Walloon and French is not whether NumP is present or not, but instead how it combines with the noun that follows. Interestingly, Bernstein argues that it combines in a way that we would not have expected from our examination of verbal syntax. She argues that the plural morpheme in Walloon begins in Num^0 and attaches to the left edge of the following noun; but, somewhat surprisingly, it shows up orthographically as the final syllable of an adjective which precedes the noun. Let's briefly examine how she arrives at this conclusion.

One fact, due to Morin (1986), that leads her in this direction is that liaison between prenominal adjectives and a following noun is absent in Walloon, though present in French.

(43) a. un gro-z -arbre
a big tree

b. une peti-t -enfant
a little child

(French)

(44) a. on gro[s] abe
a big tree

(Liège Walloon)

b. on peti[t] èfant
a small child

(Gondecourt Picard)

She suggests that the account offered here would provide an immediate explanation for this, if in Walloon there is a Num^0 that lies between the prenominal adjective and the following noun. This intervening category might then be responsible for blocking liaison in Walloon. In French, by contrast, the noun has

moved into Num^0, and is therefore in a position to trigger liaison with a preceding adjective. For this reason, she suggests that Num^0 should be present in Walloon, but not able to attract N^0s to it.

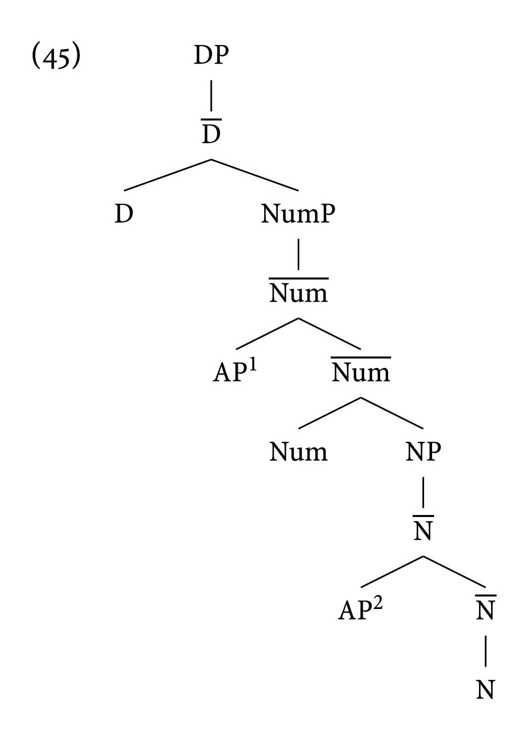
Note how this account presupposes that adjectives cannot come between Num^0 and N^0; if they could, then an intervening Num^0 cannot be blamed on the absence of liaison. This, however, is incompatible with the proposal that adjectives follow nouns (when they do) because of movement to Num^0.[9] Thus, either Bernstein must posit two differences between French and Walloon — adding that in addition to the differing expressions of noun movement they also differ in placement of adjectives — or something is being missed here. It should also be noted that this would diverge from the trend Dryer found in the relative placement of number words and adjectives. His results suggest that adjectives should be placed lower in the DP than number Num^0.

In fact, we have another kind of problem that arises as soon as we adopt the view that it is movement of Nouns to Num^0 that is responsible for the N+Adj word order. We have adopted a system that includes Earliness, which prohibits optional movement. And, as we have seen, in the Romance languages which allow the N+Adj word order, the Adj+N word order is also possible. Moreover, recall that with some adjectives, there is a difference in meaning that correlates with these two orders. We need to find a way of fitting these facts to our goal of correlating the availability of the N+Adj word order with overt movement to Num^0. One way we could do this is to imagine that adjectives can be placed either above or below Num^0, as indicated in (45) on the next page.

Once the noun has moved into Num^0, there is still, according to this model, a place for adjectives to the left of the Num^0+N^0 pair. In order to account for the meaning difference that (sometimes) arises, we might imagine that adjectives in the AP^2 position get a different interpretation (maybe restrictive) than do adjectives in the AP^1 position. We might seek an account for this difference from the fact that these adjectives are modifying different things: a Num^0 in one case and a N^0 in the other. This way of modeling the meaning difference, then, would predict that, with respect to those adjectives that show the difference, whether the adjective appears to the left or right of the noun will completely disambiguate its meaning. Thus, for example, if an adjective shows up to the left of the noun, it'll have to be non-restrictive; whereas if it appears to

9 It also runs counter the intuition that adjectives are modifying nouns, or their projections, and how we have expressed the syntactic relationship that holds between modifiers and the things they modify.

(45)



the right of the noun, it'll have to be restrictive. I don't know if this is a correct outcome.

A different way of modeling the N+Adj/Adj+N word order, that still correlates the availability of the N+Adj order with overt Noun Movement to Num^0 and would also address the problem we encountered with Bernstein's explanation for the contrast in liaison between Walloon and French, would be to hypothesize an optional projection above NumP. Then we could place adjectives on either side of this optional projection, and let the noun move into its head when it is present. I don't know what this optional phrase is, so I will simply call in "Y" here. The idea, then, would be to give to DPs the shape in (46).[10]

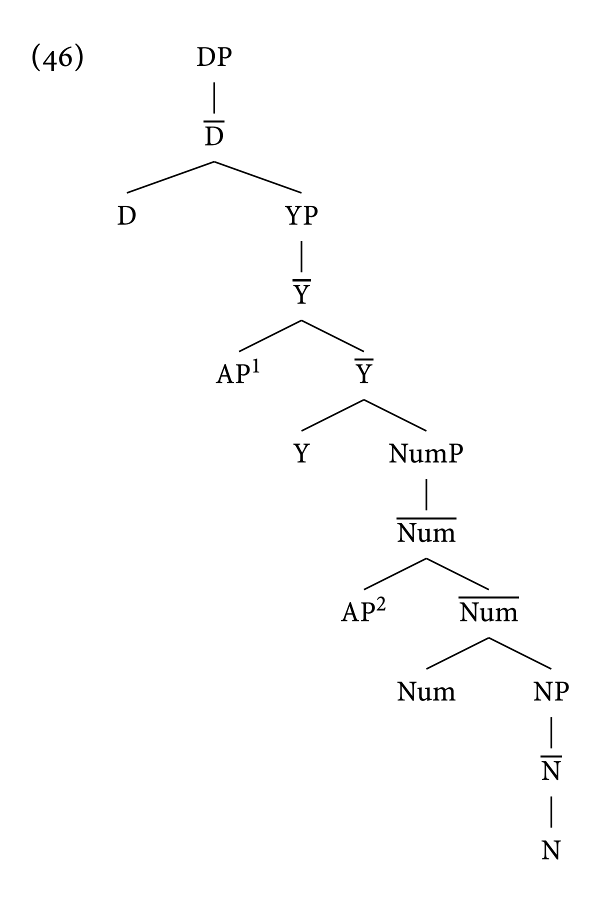
10 Relevant to this decision is that ethnic adjectives can't appear prenominally in (standard) Italian or French.

(i) * la tedesca invazione dell'Austria
the german invasion of Austria

(ii) quel tedeschissimo comportamento
that very German behavior

(Valois 1991b, p. 374)

(46)



Now, as before, let nouns be driven into Num^0 to satisfy a $*$ property associated with the number feature in Romance. When YP is absent, then nouns will surface to the right of adjectives, both those in AP^1 and those in AP^2 position. If, as before, we associate these two positions with the two interpretations that these adjectives can get, we will, in this situation, allow prenominal adjectives to have either interpretation. When YP is present, assume that Y^0 has a feature with a $*$ attribute associated with it, and the Num^0+N^0 pair will be driven into Y^0. In that situation, the noun will surface to the left of adjectives in AP^2, thus appearing to the left of adjectives with a restrictive interpretation, and will still remain to the right of adjectives in AP^2, presumably those with a non-restrictive interpretation.

To the extent, then, that ethnic adjectives show us where the D-structure position of external θ-role bearers are in nominals, this fact suggests that these external θ-role bearers are positioned before objects in Romance nominals.

Consider, by contrast, a language which does not allow Nouns to move into Num^0, presumably Walloon for instance. In these languages, movement into Y^0 will be blocked by the Head Movement Constraint. That is, if we could find something that not only prevented Nouns from surfacing in Num^0, but also prevented them from moving through Num^0, then we would have a way of correlating access to Y^0 with access to Num^0.

This, in fact, is what Bernstein does. She argues that the plural morpheme in Walloon is placed in Num^0, and this blocks movement of the noun in Walloon.

There are two plural morphemes in Walloon, one for feminine nouns and one for masculine nouns. Both are expressed orthographically on the prenominal adjective. The "feminine plural morpheme" is realized before consonant initial words as an unstressed vowel and before vowel initial words as [ɛz]. Examples follow.

(47) a. les belè[s] feyes
the pretty girls

b. dès neûrè-z -amonnes
some black berries

Compare:

(48) li neùr sipène
the black thorn

The masculine plural morpheme (-*s*) shows a similar pattern, though it is phonetically manifest only in contexts of liaison.

(49) a. dès deûr[s] tchivès
the black hair

b. dès neûr-z -ouy
the black eyes

She argues against composing the feminine plural marking of a gender morpheme and a number morpheme because this would assign to feminine the suffix -*e*, and this doesn't show up in singular nominals.

(50) li neûr sipène
the black thorn

So she supposes that there is only one morpheme, a plural one, that is to be found here. And, she conjectures that this morpheme is portmanteau with gender, or what she calls a word-marker, following work by Jim Harris. From now

on we will illustrate this morpheme with the phonetically more salient feminine one.

The evidence that these morphemes are actually attached to the noun that follows them is as follows. First, only prenominal adjectives show this morphology, as (51) indicates.

(51) a. Èle sont neûr.
they are black

b. Èle sont tot[es] petit[es].
the are very little

And when the nominal that the adjective precedes is absent, this morpheme does not appear. It's missing in copular constructions, for instance, as shown in (52).

(52) a. C'è dès bèl[es].
those are good

b. * C'è dès bèle[s].
those are good

Second, only one of these morphemes appears when two prenominal adjectives are conjoined.

(53) dès bèl[es] èt bounè[s] bièsses
some nice and good animals

This, at least, is the case in one dialect of Walloon (Boulogne Picard). In another, Liège Walloon, it is possible to find the plural morpheme on all adjectives in the series.

(54) dès bèlè[s] gradè[s] djônè[s] fèy[es]
some nice and strong young girls

She suggests that in these cases, the adjectives aren't actually stacked, but are instead conjoined. She notes that the conjunction is *es* in Walloon, and therefore homophonous with the plural morpheme.

Third, there is phonological evidence that this morpheme is a proclitic on the following noun and not suffixed onto the preceding adjective. First, there is a widespread process of final obstruent devoicing in Walloon, that Bernstein illustrates with the following pair.

(55) a. grandeûr
big

b. grande amice [grãt amis]
good friend

When adjectives are followed by the plural morpheme, they show obstruent final devoicing, as the contrast below illustrates.

(56) a. * grandè[s] fèyes
big girls

b. grantè[s] fèyes
good girls

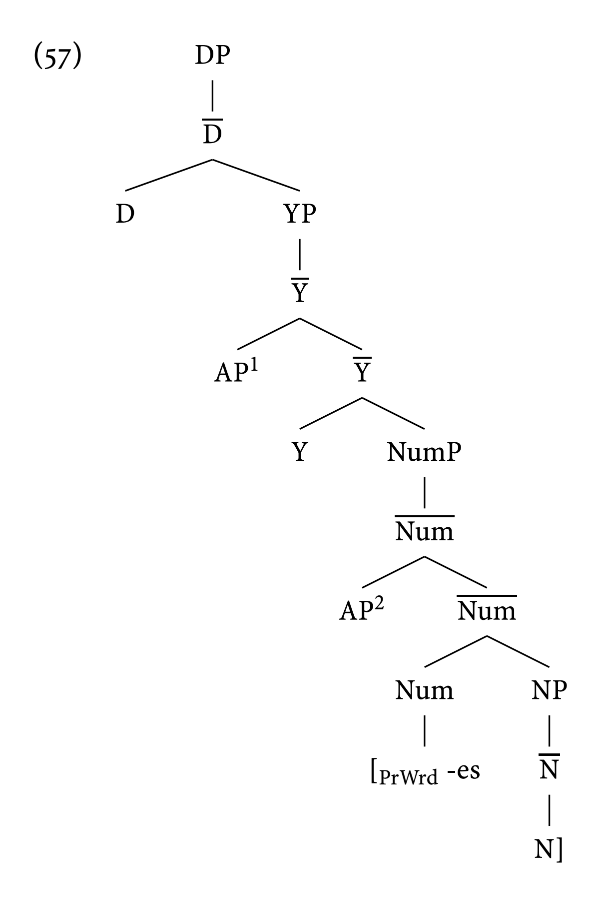
A second phonological reason for thinking that the plural affix is not part of the preceding adjective is that it is unstressed. She cites Morin who argues that all words in Walloon have stressed final syllables. Finally, again following Morin, she points to the fact that in Gondecourt Picard, the plural morpheme, *ès*, triggers harmony on the following noun. She follows Morin and adopts the proposition that harmony is restricted to words in Walloon, which leads to the conclusion that *ès* is part of the following noun, not the preceding adjective.

This pattern of data all makes sense, Bernstein concludes, if the Walloon plural suffix combines with the following noun not by way of N^0 movement, but instead, by procliticizing onto the unmoved, following N^0, as indicated in (57) on the following page. (Understand the material enclosed within "$[_{PrWd}$]" to form a phonological word.) As noted above, this will explain why Walloon nouns surface always to the right of adjectives, because they will not be able to move through Num^0 into Y^0.

Still, there are problems with this account which are central enough to suggest that it needs revision. For one thing, it persists in requiring that adjectives be placed higher than NumP, and this runs against the trend Dryer found for languages to place adjectives within NumP. In addition, it credits the availability of a noun movement past an adjective to the properties of Y^0, and only indirectly to the properties of Num^0. But the aim of Bernstein's analysis of the Walloon/French contrast is to make the properties of Num^0 responsible for noun movement past adjectives. Unless some intimate link can be made between Y^0 and Num^0, the phrase marker in (46) isn't equipped to express a correlation between occupying Num^0 and preceding single adjectives.

The decision to place adjectives higher than Num^0, and to invent a new Y^0 into which nouns can move, responds to the desire to explain the absence of liaison in Walloon between prenominal adjectives and the nouns that follow them. Bernstein's account forces Num^0 to intervene between prenominal adjectives and NP. Perhaps we should abandon trying to account for this fact,

(57)

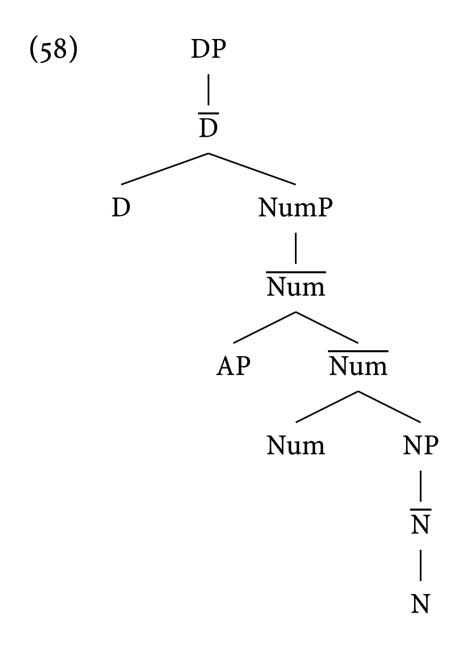


and let it come from some other idiosyncrasy of Walloon. This will allow us to return to a model of DPs like (58) on the next page.

The difference between French and Walloon, as before, consists in whether Num^0 holds a free morpheme — a clitic — as in Walloon, or a number feature that values the number feature on the following noun, as in French.

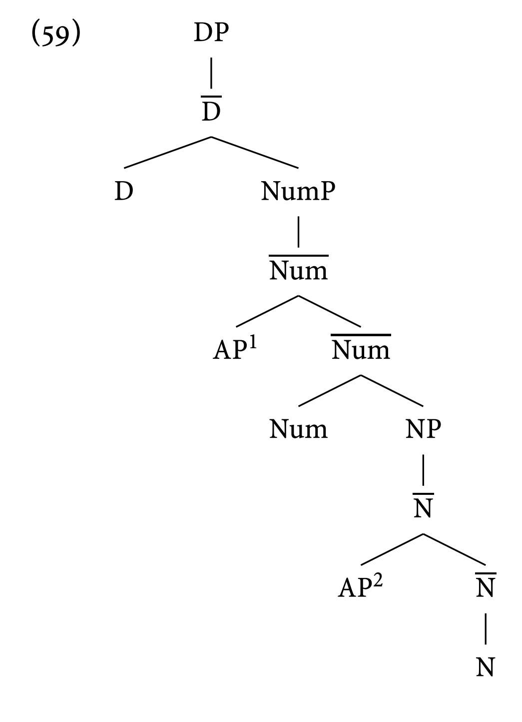
Let's consider how this account of the difference in adjective placement between Walloon and French/Spanish/Catalan might be applied to the similar difference between English and French/Spanish/Catalan. There is no evidence of the sort we've just seen in Walloon that the number morpheme in English is a free morpheme. Let's assume, then, that Num^0 in English contains a number feature, as it does in French. Given the tools developed here, perhaps the most straightforward way of modeling the difference between English and Romance, then, would be to give the number feature in French, Catalan and Spanish a $*$ property, but not give the number feature that attribute in English. This will force nouns in French, Catalan and Spanish to move to Num^0, thereby bring-

(58)

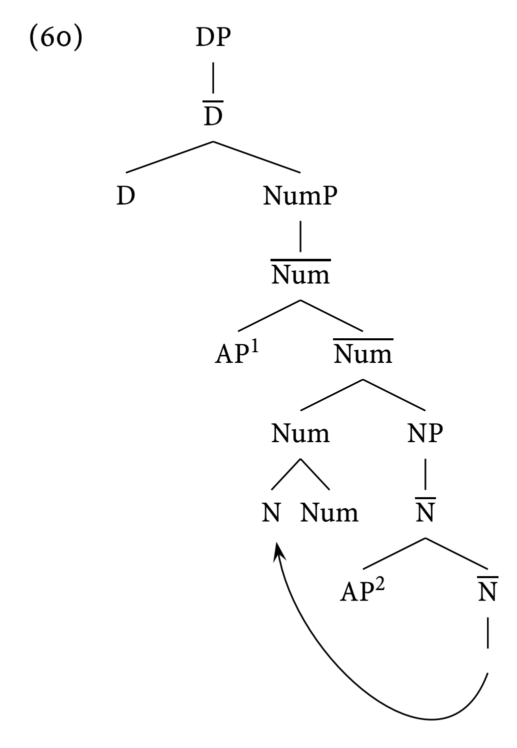


ing them to the left of (some) adjectives. So, English S-structures will arrange DPs as in (59), while in the remaining Romance languages, nouns will surface in Num^0 as in (60) below.

(59)



(60)



This gives us a three-way distinction. Walloon nouns have no number feature, English nouns do, but don't move, and French nouns have a number feature and move. The surface position of nouns in English and Walloon, then, is the same but for different reasons.

One consequence of forcing nouns in Walloon and English to remain in their underlying position is that they will remain to the right of the Specifier of NP. Recall that in Romance, we associated the ability of nouns to surface to the left of Specifier of NP with the availability of "subjects" of nouns to surface post-nominally. For instance, the French example in (16c), repeated below, arises by leaving *de chaque peintre étranger* ('of each foreign painter') in Specifier of NP and moving the noun, *portrait* ('portrait'), past it to the left.

(16c) le portrait de chaque peintre étranger de son enfant
the portrait of each painter foreign of his child
'the picture by each foreign painter of his child'

If nouns don't move to Num^0 in English or Walloon, we would expect these postnominal subjects to be unavailable in both English and Walloon. We've already seen that this is the case for English. But, interestingly, it doesn't seem to be the case for Walloon.

Walloon does allow the N+Subject word order; Bernstein illustrates this with examples like (61) below.[11] This suggests that even in Walloon, there is

(61) la fèy do mounî
the daughter of the miller
'the miller's daughter' (Bernstein 1993, (85): 241)

short noun movement, past the Specifier of NP position. If Bernstein's arguments concerning how number morphology is expressed in Walloon is correct, this short noun movement can't be to Num^0. Bernstein suggests that it is instead movement to the position associated with the "gender" morpheme that Romance nouns so typically end in. She calls this a "word marker." A schematic surface phrase marker for a Walloon DP, then, looks something like (62), then.

Note that this requires that adjectives are **not** capable of being within NP. Indeed, Walloon illustrates that the availability of postnominal subjects and postnominal (single) adjectives do not correlate cross-linguistically. It is necessary, therefore, to divorce the processes that yield these two word-orders, and if noun movement is the relevant process, then this means there must be two positions to which nouns can move, with subjects below, and adjectives above, the lower of these. Up to now we have relied on a general theory of modifier placement one of whose outcomes is that adjectives should be adjoined to the $\overline{N}$ that they modify. One thing we learn from this study, then, is that this general theory will have to be changed.

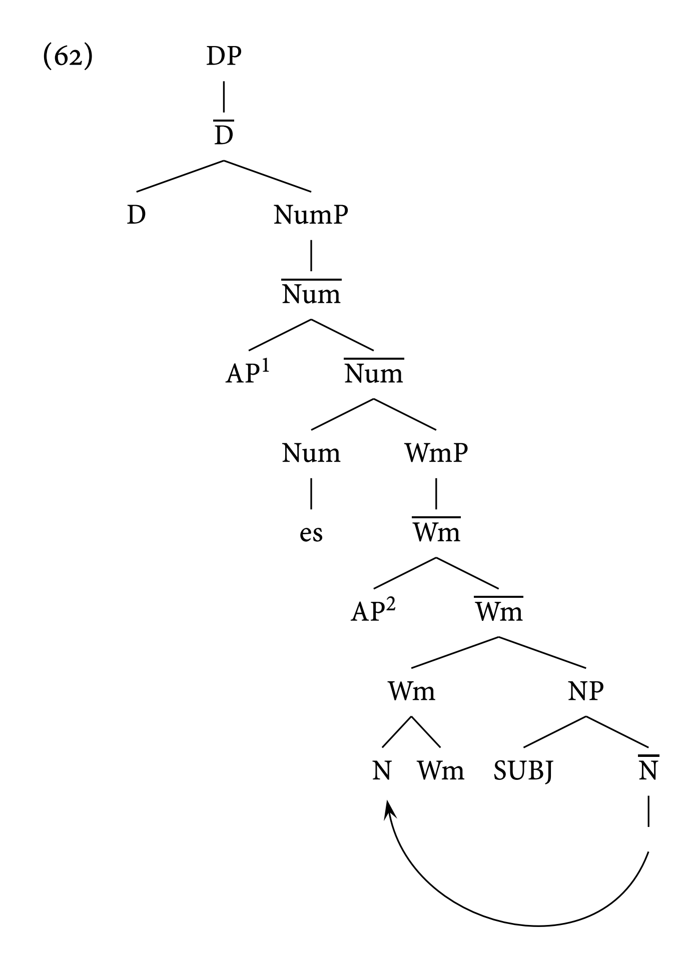
But let's leave that for a later occasion.

We must also revisit our account for why postnominal subjects don't arise in English. It's no longer sufficient to prevent nouns from moving to Num^0 in English. We must also now address the possibility that nouns could move to a position beneath Num^0: the Wm^0 position that Bernstein posits Wallon to have, for instance.

While it might be that there are no postnominal subjects in English because English nouns don't make even a short move, but it could also be because

11 This is perhaps not the most compelling example as it is difficult to tell whether *miller* bears the "subject" relation to *daughter*. Interestingly, Bernstein claims that Walloon also allows for postnominal adjectives when they are ethnic; in fact, in these situations, the prenominal position is blocked. This also, rather dramatically, supports the conclusion that "subjects" can be postnominal in Walloon.

(62)



the other ingredient necessary to getting postnominal subjects is missing from English. Perhaps subjects cannot remain in Specifier of NP position. We might characterize this difference between English and Romance in terms of the positions that Case is assigned to within DPs. Let the Specifier of DP be assigned Case in both English and Romance, but let only Romance assign Case to Specifier of NP. Note that this Case is manifest in what appears to be a preposition — *di* or *de*, depending on the language. Let's call this Case, the one expressed by a preposition, "Nominal Case." On this view, then, the difference in availability of postnominal subjects between English and Romance boils down to the availability of Nominal Case in Specifier of NP.

Indeed, the "subject" arguments of DPs uniformly appear with the genitive Case in English, and this is a position, as we've seen, associated Specifier of DP. Thus, no matter what its position relative to the noun, the subject of a "transitive" noun cannot be Case marked with *of*, as (63) indicates.

(63) a. * the discussion of Jill of the problem
b. * the discussion of the problem of Jill
c. * the of Jill discussion of the problem
d. * the placement of Mark of the sofa
e. * the placement of the sofa of Mark
f. * the of Mark placement of the sofa

It's not possible to express the subjects of nouns this way even when the nouns do not express their object argument. Leaving the objects unexpressed in the examples in (63), for example, does not improve them.

(64) a. * the discussion of Jill
b. * the of Jill discussion
c. * the placement of Mark
d. * the of Mark placement

But it's not that Nominal Case is completely absent in English. It is possible for Nominal Case to be found on the arguments of nouns that derive from unaccusative verbs, as in (65).

(65) a. the death of her
b. the arrival of it
c. the appearance of Jill
d. the sinking of the ship

With nouns derived from unergative verbs, the situation is somewhat intermediate, as illustrated by (66) below.[12] If we interpret these facts as indicating that

(66) a. ?* the running of her
b. * the talking of him
c. ?? the dancing of Jill
d. ?* the speaking of the woman
e. ?* the sitting of Mark

there is a distinction between the "unaccusative" nouns and the others — that

12 See Grimshaw (1990) for a discussion of these facts and an argument that nouns divide into the unaccusative and unergative classes.

is if we set aside the cause for the intermediate status of the "unergative" nouns — then this pattern can be described with (67).

(67) Nominal Case Assignment: English
Nominal Case is assigned to the "object position" of nouns.

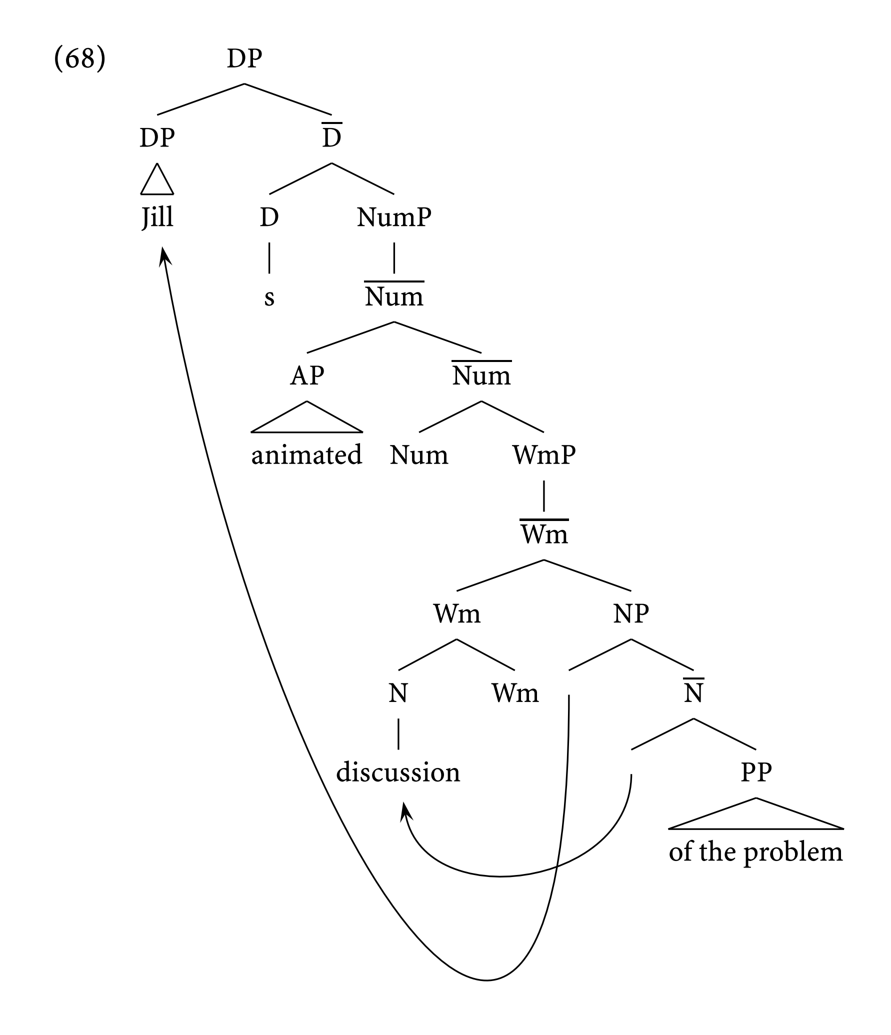
We'll set to defining what "object position" means later; but, importantly, it can't have the same sort of definition we've given to the positions that verbs assign their "object" Case to if we adopt the view that nouns move overtly to Wm^0 in English. Object Case is assigned by verbs to positions they govern (and are adjacent to), as we've seen. If we let Nominal Case be assigned by nouns to positions they govern, then once a noun has moved to Wm^0 it should be able to assign its Case to a DP within Specifier of NP: just the effect we are hoping to avoid.

It would be reasonable, therefore, to expect the general absence of postnominal subjects in English DPs to be caused by the constraints on Nominal Case that derive (67). This means it is conceivable that nouns in English do make a short movement, as they do in Walloon. This hypothesis, then, would give an example like "Jill's animated discussions of the problem" a representation like that in (68) on the next page.

As this discussion makes clear, the relative heights of Wm^0 and Num^0 correlates the relative position of nouns and subjects with the relative position of nouns and adjectives. The general prediction is that there should be a positive correlation between nouns surfacing to the left of (bare) adjectives and nouns surfacing to the left of subjects. We should not find[13] a language, in other words, that is the anti-Walloon: nouns surface to the left of bare adjectives but cannot surface to the left of "subjects." In fact, this correlation does seem to hold in our language sample. All the languages we have examined that allow nouns to surface to the left of adjectives also allow them to surface to the left of subjects.

13 Restricting attention to just those languages that Case mark subject DPs in Specifier of NP, and have the ban on right-adjoining bare adjectives.

(68)


DP
DP
D̄
Jill
D
NumP
s
Num̄
AP
Num̄
animated
Num
WmP
Wm̄
Wm
NP
N
Wm
N̄
discussion
PP
of the problem

Bibliography

Abney, Steven. 1987. The English noun phrase in its sentential aspect. Doctoral Dissertation, Massachusetts Institute of Technology.

Adger, David. 2003. *Core syntax*. Oxford: Oxford University Press.

Alexiadou, Artemis, and Elena Anagnostopoulou. 1998. Parametrizing Agr: word order, V-movement and EPP-checking. *Natural Language and Linguistic Theory* 16:491–539.

Alexiadou, Artemis, and Elena Anagnostopoulou. 1999. EPP without Spec, IP. In *Specifiers*, ed. David Adger, Bernadette Plunkett, George Tsoulas, and Susan Pintzuk, 93–109. Oxford: Oxford University Press.

Baker, C. L. 1978. *Introduction to generative transformational syntax*. Englewood Cliffs, New Jersey: Prentice-Hall.

Baker, Mark. 1985. The mirror principle and morphosyntactic explanation. *Linguistic Inquiry* 16:373–416.

Baker, Mark C. 2003. *Lexical categories*. Cambridge: Cambridge University Press.

Baltin, Mark. 1982. A landing site theory of movement rules. *Linguistic Inquiry* 13:1–38.

Belletti, Adriana. 1982. On the anaphoric status of the reciprocal construction in Italian. *The Linguistic Review* 2:101–138.

Belletti, Adriana. 1990. *Generalized verb movement*. Rosenberg & Sellier.

Bernstein, Judy. 1991. DPs in French and Walloon: Evidence for parametric variation in nominal head movement. *Probus* 3:101–126.

Bernstein, Judy B. 1993. Topics in the syntax of nominal structure across Romance. Doctoral Dissertation, City University of New York.

den Besten, Hans. 1983. On the interaction of root transformations and lexical deletive rules. In *On the formal syntax of the Westgermania*, ed. W. Abraham. Amsterdam, The Netherlands: John Benjamins Publishing Company.

Bobaljik, Jonathan. 1994. What does adjacency do? In *The morphology-syntax connection*, ed. Heidi Harley and Colin Phillips, 1–32. MIT.

Bobaljik, Jonathan. 1995. Morphosyntax: The syntax of verbal inflection. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge.

Bobaljik, Jonathan. 2003. Floating quantifiers: Handle with care. In *The second Glot International state-of-the-article book*, ed. Lisa L.-S. Cheng and Rint Sybesma, 107–148. Mouton de Gruyter.

Bod, Rens, Jennifer Hay, and Stefanie Jannedy. 2003. *Probabilistic linguistics*. Cambridge, Massachusetts: MIT Press.

Boeckx, Cedric, and Norbert Hornstein. 2003. Reply to "Control is not movement". *Linguistic Inquiry* 34:269–280.

Boeckx, Cedric, and Norbert Hornstein. 2004. Movement under control. *Linguistic Inquiry* 35:431–452.

Boeckx, Cedric, and Norbert Hornstein. 2006a. Control in Icelandic and theories of control. *Linguistic Inquiry* 37:591–606.

Boeckx, Cedric, and Norbert Hornstein. 2006b. The virtues of control as movement. *Syntax* 9:118–130.

Boersma, Paul, and Bruce Hayes. 2001. Empirical tests of the gradual learning algorithm. *Linguistic Inquiry* 32:45–86.

Bošković, Željko. 2001. Floating quantifiers and theta-role assignment. In *North East Linguistic Society*, ed. Minjoo Kim and Uri Strauss, 59–78. Georgetown University: GLSA.

Bošković, Željko. 2002a. A-movement and the EPP. *Syntax* 5:167–218.

Bošković, Željko. 2002b. Expletives don't move. In *Proceedings of North East Linguistic Society 32*, ed. Masako Hirotani, volume 1, 21–40.

Bošković, Željko. 2004. Be careful where you float your quantifiers. *Natural Language and Linguistic Theory* 22:681–742.

Bowers, John. 1993. The syntax of predication. *Linguistic Inquiry* 24:591–656.

Bresnan, Joan. 1972. Theory of complementation in English syntax. Doctoral Dissertation, Massachusetts Institute of Technology.

Burzio, Luigi. 1986. *Italian syntax*. Reidel Publishers.

Carrier, Jill, and Janet H. Randall. 1992. The argument structure and syntactic structure of resultatives. *Linguistic Inquiry* 23:173–234.

Chierchia, Gennaro. 1984. Topics in the syntax and semantics of infinitives and gerunds. Doctoral Dissertation, University of Massachusetts, Amherst.

Chomsky, Noam. 1957. *Syntactic structures*. Mouton.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Massachusetts: M.I.T. Press.

Chomsky, Noam. 1970. Remarks on nominalization. In *Readings in English transformational grammar*, ed. J. Jacobs and P. Rosenbaum, 184–221. Waltham, Massachusetts: Ginn.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht, The Netherlands: Foris Publications.

Chomsky, Noam. 1985. *The logical structure of linguistic theory*. Chicago: University of Chicago Press.

Chomsky, Noam. 1986a. *Barriers*. Cambridge, Massachusetts: MIT Press.

Chomsky, Noam. 1986b. *Knowledge of language*. New York, New York: Praeger Publishers.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, Massachusetts: MIT Press.

Chomsky, Noam. 1998. Some observations on economy in generative grammar. In *Is the best good enough?*, ed. Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis, and David Pesetsky, 115–128. Cambridge, Massachusetts: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. Roger Martin, David Michaels, and Juan Uriagereka, 89–156. MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in linguistics*, ed. Michael Kenstowicz, 1–52. Cambridge, Massachusetts: MIT Press.

Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of English*. New York: Haper & Row, Publishers.

Chomsky, Noam, and Howard Lasnik. 1993. Principles and parameter theory. In *Syntax: an international handbook of contemporary research*, ed. Arnim von Stechow, W. Sternefeld, and T. Vennemann, 506–569. Berlin: Walter de Gruyter.

Chomsky, Noam, and George Miller. 1963. Finitary models of language users. In *Handbook of Mathematical Psychology 2*, ed. Luce, Bush, and Galanter, volume Handbook of Mathematical Psychology 2, 419–491. New York, New York: Wiley and Sons.

Déprez, Vivian. 1989. On the typology of syntactic positions and the nature of chains. Doctoral Dissertation, MIT, Cambridge, Massachusetts.

Diesing, Molly. 1992. *Indefinites*. Cambridge, Massachusetts: MIT Press.

Doetjes, Jenny. 1995. Quantification at a distance and iteration. In *Proceedings of the North East Linguistic Society 25*, ed. Jill Beckman, 111–126. University of Pennsylvania: Graduate Linguistic Student Association.

Dryer, Matthew S. 1989. Plural words. *Linguistics* 27:865–895.

Dryer, Matthew S. 1992. The Greenbergian word order correlations. *Language* 68:81–138.

Emonds, Joseph E. 1976. *A transformational approach to English syntax*. New York, New York: Academic Press.

Epstein, Samuel David, Acrisio Pires, and T. Daniel Seely. 2005. EPP in T: More controversial subjects. *Syntax* 8:65–80.

Fillmore, Charles. 1968. The case for case. In *Universals in linguistic theory*, ed. Emmon Bach and Richard Harms, 1–90. New York, New York: Holt, Rinehart and Winston.

Fox, Danny. 2000. *Economy and semantic interpretation*. Cambridge, Massachusetts: MIT Press.

Frazier, Lyn. 1978. On comprehending sentences: Syntactic parsing strategies. Doctoral Dissertation, University of Connecticut, Storrs.

Fukui, Naoki, and Margaret Speas. 1986. Specifiers and projection. In *MIT working papers in linguistics*, ed. Naoki Fukui, Tova R. Rapoport, and Elizabeth Sagey, volume 8, 128–172. Department of Linguistics and Philosophy.

Gerdts, Donna. 1981. Object and absolutive in Halkomelem Salish. Doctoral Dissertation, University of California, San Diego.

Gibson, Edward. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68:1–76.

Gibson, Jean. 1980. Clause union in Chamorro and in universal grammar. Doctoral Dissertation, University of California, San Diego.

Giorgi, Alessandra, and Giuseppe Longobardi. 1991. *The syntax of noun phrases*. Cambridge: Cambridge University Press.

Giusti, Giuliana. 1990a. Floating quantifiers in Germanic. In *Grammar in progress*, ed. Joan Mascaró and Marina Nespor, 137–146. Dordrecht, The Netherlands: Foris Publications.

Giusti, Giuliana. 1990b. Floating quantifiers, scrambling and configurationality. *Linguistic Inquiry* 21:633–641.

Godfrey, J., E. Holliman, and J. McDaniel. 1992. SWITCHBOARD: telephone speech corpus for research and development. In *Proceedings of ICASSP 92*, 517–520. San Francisco.

Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In *Universals of language*, ed. Joseph Greenberg. Cambridge, Massachusetts: MIT Press.

Grimshaw, Jane. 1979. Complement selection and the lexicon. *Linguistic Inquiry* 10:279–326.

Grimshaw, Jane. 1990. *Argument structure*. Cambridge, Massachusetts: MIT Press.

Groat, Erich. 1999. Raising the case of expletives. In *Working minimalism*, ed. Samuel Epstein and Norbert Hornstein, 27–43. Cambridge, Massachusetts: MIT Press.

Groat, Erich M. 1995. English expletives: A minimalist approach. *Linguistic Inquiry* 26:354–365.

Gruber, Jeffrey S. 1965. Studies in lexical relations. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Hackl, Martin. 2000. Comparative quantifiers. Doctoral Dissertation, Massachusetts Institute of Technology.

Harris, Zellig. 1946. From morpheme to utterance. *Language* 22:161–183.

Hazout, Ilan. 2004. The syntax of existential constructions. *Linguistic Inquiry* 35:393–430.

Heim, Irene. 1987. Where does the definiteness restriction apply evidence from the definiteness of variables. In *The representation of (in)definiteness(14)*, ed. Eric J. Reuland and Alice G. B. ter Meulen, 21–42. Cambridge, Massachusetts: MIT Press.

Hornstein, Norbert. 1999. Movement and control. *Linguistic Inquiry* 30:69–96.

Hornstein, Norbert. 2000. Control in GB and minimalism. In *The first glot international state-of-the-article book*, ed. Lisa L.-S. Cheng and Rint Sybesma, 27–45. New York: Mouton de Gruyter.

Huang, C.-T. James. 1982. Logical relations in Chinese and the theory of grammar. Doctoral Dissertation, Massachusetts Institute of Technology.

von Humboldt, Wilhelm. 1836. *über die Verschiedenheit des Menschlichen Sprachbaues und Ihren Einfluss auf die Geistige Entwickelung des Menschengeschlechts*. Berlin: Royal Academy of Sciences of Berlin.

Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar*. Cambridge, Massachusetts: MIT Press.

Jackendoff, Ray. 1977. *X′ syntax*. Cambridge, Massachusetts: MIT Press.

Jaeggli, Osvaldo. 1982. *Topics in Romance syntax*. Dordrecht Holland: Foris Publications.

Kayne, Richard S. 1975. *French syntax: the transformational cycle*. Cambridge, Massachusetts: MIT Press.

Kayne, Richard S. 1983. Connectedness. *Linguistic Inquiry* 14:223–249.

Kayne, Richard S. 1984. *Connectedness and binary branching*. Dordrecht, Holland: Foris Publications.

Kimball, J. 1973. Seven principles of surface structure parsing in natural language. *Cognition* 2:15–47.

Kitagawa, Yoshihisa. 1986. Subjects in Japanese and English. Doctoral Dissertation, University of Massachusetts, Amherst.

Klima, Edward S. 1964. Negation in English. In *The structure of language: Readings in the philosophy of language*, ed. Jerry A. Fodor and Jerrold J. Katz, 246–323. Prentice Hall.

Koopman, Hilda, and Dominique Sportiche. 1991. The position of subjects. *Lingua* 85:211–258.

Koster, Jan. 1975. Dutch as an SOV language. *Linguistic Analysis* 1:111–136.

Koster, Jan. 1978. Why subject sentences don't exist. In *Recent transformational studies in European languages*, ed. Samuel Jay Keyser. Cambridge, Massachusetts: MIT Press.

Kratzer, Angelika. 2005. Building resultatives. In *Event arguments in syntax, semantics, and discourse*, ed. Claudia Maienborn and Angelika Wöllstein-Leisten, 177–212. Tübingen: Niemeyer.

Kuroda, S.-Y. 1988. Whether we agree or not. *Linguisticae Investigationes* 12:1–47.

Landau, Idan. 2004. The scale of finiteness and the calculus of control. *Natural Language and Linguistic Theory* 22:811–877.

Langacker, Ronald W. 1969. On pronominalization and the chain of command. In *Modern studies in English*, ed. David A. Reibel and Sandford A. Schane, 160–186. Prentice Hall.

Lasnik, Howard. 1981. Restricting the theory of transformations. In *Explanation in linguistics*, ed. Norbert Hornstein and David Lightfoot. London: Longmans.

Lasnik, Howard. 1992. Case and expletives: Notes toward a parametric account. *Linguistic Inquiry* 23:381–406.

Lasnik, Howard. 1995. Case and expletives revisited: On greed and other human failings. *Linguistic Inquiry* 26:615–634.

Lasnik, Howard, and J. Kupin. 1977. A restrictive theory of transformational grammar. *Theoretical Linguistics* 4:173–196.

Lasnik, Howard, and Mamoru Saito. 1984. On the nature of proper government. *Linguistic Inquiry* 15:235–290.

Levin, Beth. 1993. *English verb classes and alternations*. Chicago: University of Chicago Press.

Levin, Beth, and Malka Rappaport. 1986. The formation of adjectival passives. *Linguistic Inquiry* 17:623–661.

Marantz, Alec. 1984. *On the nature of grammatical relations*. Cambridge, Massachusetts: MIT Press.

Marcus, Mitchell, Beatrice Santorini, and May Ann Marcinkiewicz. 1993. Building a large annotated corpus of English: The Penn treebank. In *Computational Linguistics 19*, 313–330.

Martin, Roger. 1992. On the distribution and case features of PRO. Doctoral Dissertation, University of Connecticut, Storrs, University of Connecticut, Storrs.

Martin, Roger. 2001. Null case and the distribution of PRO. *Linguistic Inquiry* 32:141–166.

McCawley, James D. 1970. English as a VSO language. *Language* 46:286–299.

Milsark, Gary. 1974. Existential sentences in English. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Miyagawa, Shigeru. 1989. *Structure and case marking in Japanese*. San Diego: Academic Press.

Moro, Andrea. 1997. *The raising of predicates : predicative noun phrases and the theory of clause structure*. Cambridge ; New York ;: Cambridge University Press.

Muysken, Pieter. 1979. Quechua causatives and Logical Form: A case study in markedness. In *Theory of markedness in generative grammar*, ed. Adriana Belletti, Luigi Brandi, and Luigi Rizzi. Scuola Normale Superiore, Pisa.

Muysken, Pieter. 1981. Quechua word structure. In *Binding and filtering*, ed. Frank Heny. Cambridge, Massachusetts: MIT Press.

Partee, Barbara Hall, Alice G. B. ter Meulen, and Robert Eugene Wall. 1990. *Mathematical methods in linguistics*. Dordrecht ; Boston: Kluwer Academic.

Perlmutter, David. 1971. *Deep and surface structure constraints in syntax*. New York: Holt, Rinehart and Winston.

Perlmutter, David. 1978. Impersonal passives and the unaccusative hypothesis. In *Berkeley Linguistic Society*, volume IV, 157–189. University of California, Berkeley.

Pesetsky, David. 1982. Paths and categories. Doctoral Dissertation, Massachusetts Institute of Technology.

Pesetsky, David. 1989. The earliness principle. Paper presented at GLOW.

Pires, Acrisio. 2001. Clausal and TP-defective gerunds: control without tense. In *North East Linguistic Society*, ed. Minjoo Kim and Uri Strauss, 390–406. Georgetown University: GLSA.

Pires, Acrisio. 2006. *The Minimalist syntax of Defective Domains*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Pollock, Jean-Yves. 1989. Verb movement, UG and the structure of IP. *Linguistic Inquiry* 20:365–424.

Postal, Paul M. 1969. On so-called 'pronouns' in English. In *Modern studies in English*, ed. D. Reibel and Sandford Schane, 201–244. Englewood Cliffs, New Jersey: Prentice-Hall.

Reinhart, Tanya. 1976. The syntactic domain of anaphora. Doctoral Dissertation, Massachusetts Institute of Technology.

Rizzi, Luigi. 1982. *Issues in Italian syntax*. Dordrecht, Holland: Foris Publications.

Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge, Massachusetts: MIT Press.

Roberts, Craige. 1987. Modal subordination, anaphora and distributivity. Doctoral Dissertation, University of Massachusetts, Amherst.

Rochemont, Michael, and Peter W. Culicover. 1990. *English focus constructions and the theory of grammar*. Cambridge: Cambridge University Press.

Ross, John. 1967. Constraints on variables in syntax. Doctoral Dissertation, Massachusetts Institute of Technology.

Roussou, Anna. 2001. Control and raising in and out of subjunctive complements. In *Comparative syntax of Balkan languages*, ed. María Luisa Rivero and Angela Ralli, 74–104. Oxford: Oxford University Press.

Safir, Kenneth. 1985. *Syntactic chains*. Cambridge: Cambridge University Press.

Safir, Kenneth J. 1987. What explains the definiteness effect? In *The representation of (in)definiteness(14)*, ed. Eric J. Reuland and Alice G. B. ter Meulen, 71–97. Cambridge, Massachusetts: MIT Press.

Saito, Mamoru. 1985. Some asymmetries in Japanese and their theoretical implications. Doctoral Dissertation, Massachusetts Institute of Technology.

Saito, Mamoru, and Naoki Fukui. 1998. Order in phrase structure and movement. *Linguistic Inquiry* 29:439–474.

Sapir, Edward. 1921. *Language: an introduction to the study of speech*. New York: Harcourt, Brace and Co.

Sigurðsson, Halldór Ármann. 1991. Icelandic case-marked PRO and the licensing of lexical arguments. *Natural Language and Linguistic Theory* 9:327–363.

Sportiche, Dominique. 1988. A theory of floating quantifiers and its corollaries for constituent structure. *Linguistic Inquiry* 19:425–449.

Stechow, Arnim von. 1996. Against LF pied-piping. *Natural Language Semantics* 4:57–110.

Stowell, Tim. 1978. What was there before there was there. In *Proceedings of the Fourteenth Regional Meeting of the Chicago Linguistic Society*, ed. D. Farkas, volume Proceedings of the Fourteenth Regional Meeting of the Chicago Linguistic Society. Chicago, Illinois: The University of Chicago.

Stowell, Timothy. 1981. Origins of phrase structure. Doctoral Dissertation, Massachusetts Institute of Technology.

Terzi, Arhonto. 1992. PRO and obviation in modern Greek subjunctives. In *The Proceedings of the Tenth West Coast Conference on Formal Linguistics*, ed. Dawn Bates, 471–482. Stanford, California: Center for the Study of Language and Information.

Travis, Lisa. 1984. Parameters and effects of word order variation. Doctoral Dissertation, Massachusetts Institute of Technology.

Valois, Daniel. 1991a. The internal syntax of DP. Doctoral Dissertation, University of California, Los Angeles.

Valois, Daniel. 1991b. The internal syntax of DP and Adjective placement in French and English. In *The Proceedings of the North Eastern Linguistics Society*, ed. Tim Sherer, volume The Proceedings of the North Eastern Linguistics Society, 367–382. Universite du Quebec a Montreal: The Graduate Students Linguistic Association, The University of Massachusetts-Amherst.

Vikner, Sten. 1995. *Verb movement and expletive subjects in the Germanic languages*. Oxford: Oxford University Press.

Wasow, Tom. 1977. Transformations and the lexicon. In *Formal syntax*, ed. Peter Culicover, Tom Wasow, and Adrian Akmajian, 327–360. New York, New York: Academic Press.

Williams, Edwin. 1980. Predication. *Linguistic Inquiry* 11:203–238.

Williams, Edwin. 1989. The anaphoric nature of theta roles. *Linguistic Inquiry* 20:425–456.